

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 03:15:26 : Search time 101 Seconds
(without alignments)
5902.769 Million cell updates/sec

Title: US-10-086-464-1
Perfect score: 1444
Sequence: 1 atgtctccggcgccgtctcc.....atagtggacctctcttttaa 1944

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 44362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	9.4	2868	US-09-228-986-4	Sequence 4, Appli
2	178.4	9.2	3239	US-09-228-986-9	Sequence 9, Appli
3	174.6	9.0	2432	US-09-228-986-7	Sequence 7, Appli
4	166.4	8.6	2749	US-08-265-628-1	Sequence 1, Appli
5	162.4	8.4	4104	US-08-881-708-1	Sequence 1, Appli
6	146.6	7.5	2571	US-07-717-331F-9	Sequence 9, Appli
7	146.6	7.5	2833	US-07-717-331F-1	Sequence 1, Appli
8	140.6	7.2	1554	US-08-587-680A-24	Sequence 24, Appli
9	139.8	7.2	2749	US-07-717-331F-4	Sequence 4, Appli
10	138.2	7.1	2389	US-09-228-986-1	Sequence 1, Appli
11	130.8	6.7	2336	US-09-228-986-10	Sequence 10, Appli
12	121.2	6.2	966	US-08-447-185-2	Sequence 2, Appli
13	121.2	6.2	2443	US-08-447-185-3	Sequence 3, Appli
14	117.8	6.1	3097	US-09-228-986-11	Sequence 11, Appli
15	117.8	6.1	16442	US-08-781-891-208	Sequence 208, Appli
16	115.2	5.9	1926	US-09-249-585A-2	Sequence 2, Appli
17	115.2	5.9	2580	US-09-050-863-2	Sequence 2, Appli
18	115.2	5.9	2580	US-09-359-081-2	Sequence 2, Appli
19	115.2	5.9	5452	US-09-130-114-1	Sequence 1, Appli
20	115.2	5.9	9600	US-08-910-647-1	Sequence 1, Appli
21	115.2	5.9	9600	US-09-620-925-1	Sequence 1, Appli
22	115.2	5.9	10596	US-07-884-811-15	Sequence 15, Appli
23	115.2	5.9	10596	US-07-885-971-15	Sequence 15, Appli
24	115.2	5.9	10596	US-08-087-783A-15	Sequence 15, Appli
25	115.2	5.9	10596	US-08-194-088B-15	Sequence 15, Appli
26	115.2	5.9	10596	US-08-194-087-15	Sequence 15, Appli
27	115.2	5.9	10596	PCT-US93-04648-15	Sequence 15, Appli

28 114.4 5.9 2943 4 US-09-503-922-2 Sequence 2, Appli
29 110.8 5.7 7218 1 US-08-232-463-14 Sequence 14, Appli
30 107.6 5.5 51259 3 US-08-781-891-209 Sequence 209, Appli
31 101.2 5.2 5733 2 US-08-473-553A-1 Sequence 1, Appli
32 100.2 5.2 1505 1 US-07-915-246-1 Sequence 1, Appli
33 98.8 5.1 2649 4 US-09-238-986-12 Sequence 12, Appli
34 98.2 5.1 913 1 US-08-217-327-3 Sequence 3, Appli
35 98.2 5.1 913 1 US-07-385-970A-3 Sequence 3, Appli
36 98.2 5.1 913 1 US-08-298-687A-3 Sequence 3, Appli
37 98.2 5.1 913 1 US-08-530-797-2 Sequence 2, Appli
38 98.2 5.1 913 1 US-08-298-829-3 Sequence 3, Appli
39 98.2 5.1 913 2 US-08-787-335-2 Sequence 2, Appli
40 96.6 5.0 1984 1 US-07-885-970A-25 Sequence 25, Appli
41 96.6 5.0 1985 1 US-08-298-687A-25 Sequence 25, Appli
42 96.6 5.0 1985 1 US-08-298-829-25 Sequence 25, Appli
43 96.6 5.0 3489 2 US-08-728-323A-1 Sequence 1, Appli
44 96.6 5.0 3489 4 US-09-298-568-1 Sequence 1, Appli
45 96.6 5.0 32207 2 US-08-770-379-20 Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-09-228-986-4
; Sequence 4, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 2868
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-4

Query Match 9.4%; Score 182; DB 4; Length 2868;
Best Local Similarity 52.9%; Pred. No. 8.9e-36;
Matches 467; Conservative 0; Mismatches 400; Indels 15; Gaps 3;
QY 787 TTCACATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGGCGAACTTGTAGGA 846
DB 1116 TTCTCATTCGGGAATTACAGGTTGCACGTGATGGTTTACCAATAGAACATCTCTGGC 1175
QY 847 CAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTGCTAGTGGGAAGAAGTTCTGTG 906
DB 1176 AGAGGTGGTTTGGAAAGGTGTACAAAGGCGCGCTTCAGATGTTCTCTGGTGGCTGTA 1235
QY 907 AACCACTTGAAGTTGGAGTGTCTGAG---GGAGAGGAGGAGTTTACGGCAGAGGTTGAG 963
DB 1236 AAACGCTCTGAAGGAAGAGCGTACCCGGTGGAGAGTTGCAGTTTCAACAGAAAGTGGAG 1295
QY 964 ATCATCAGCAGAGTTTCCACACAGGCTCTGGTGTCTCTTGTGTTATTGCATCCCGGT 1023
DB 1296 ATGATGAGCATGCACTACATAGAACCTCTTCGAGTACGTGGATCTGCATGACACCC 1355
QY 1024 GCCAAGATGCTTCTCTATGAGTTTGTTCCTAACCAACATCTCGAGCTTCACTCCAT 1383
DB 1356 ACTGAACGGGTGTTTGTTCCTCCCTACATGGCAATGGAAGTGTGTTCTCATGCTACGA 1415
QY 1084 GCGGAGGAC-----GGCCTACATGGAATGAGCAGCACCAGATGAGATGCTCTTGA 1137
DB 1416 GAGAGGACAAATGACCCACCTTAGATTGGCAACTCGCAAGCGCATAGCATGGGT 1475
QY 1138 TCTGCTAAAGGACTTCTTATCTTCATGAGATTGCAATCTTAATCAATCATTCACCGAT 1197

Db 1476 TGTGGAAGAGGCTCTCTCTACTTGTGATGA...CATGTGTGATCTTAAAGATTATTCACCGGAT 1535
QY 1198 ATCAAGGCTTCAACATATGATAGATTCAGT...TGAAGCTAAGGTTGCGATTTGGT 1257
Db 1536 GTCAAGGCTGCAACATCTTACTTGGATGAAG...ATATGAGGCACTGTGGGGGATTTGGC 1595
QY 1258 CTGTCTAAGATTCCTTCTGATACAAACAGG...CATGTATCAACAGCTGTGTGGGAACCTTT 1317
Db 1596 TTGGCAAACTTATGATATATAGGACACACAT...GTGTACGAGGCTGTCTGTGGAACCATI 1655
QY 1318 GGTACTTGGCTCCGGAATACGCTGCAAGGG...AAAGCTCACGGAGAGTCTGAGTTTC 1377
Db 1656 GGCACATAGCAGCTGAGTACCTTTCTACTG...GAAGCTCTTCGGAAGAGACAGAGCTATT 1715
QY 1378 TCATTGGGCTTGTCTTTTGGAGCTCATTA...CTACTGAGAGCTGCGACCGTTGATGCCAAT 1437
Db 1716 GATATGGATCAATGCTGGAGCTTATACGG...GACAGGCAATTTG-----ACCTT 1769
QY 1438 GTCTATGTAGACAGCTTATGTAGTGGG...GACAGCAATTTGTTAAACGAGCATCTGAG 1497
Db 1770 GCAGCTTTACAAATGATGATGATGTCATG...TCTTGCCTTACGTTAAAGGCTTACTAAA 1829
QY 1498 CAGGAGACTTTGAGGCTTGTAGCTGATG...CAAGATGATTAATGGTATGACAGAGAG 1557
Db 1830 GAGAGAGGCTTGATATGCTAGTGTGAT...CTTAAAGAACATTAATGTGTGAAGCAG 1889
QY 1558 ATGGCTCGCATGCTTGTCTGCTGCGCTG...TGTGTCGCCATTCAGCTCGCGCAGACCT 1617
Db 1890 GTGGACACTTATCAAGTTGCTATTTACT...TTGTACACAGGCTCACCAATGGATAGCCA 1949
QY 1618 CGCATGAGCAGATTTGCGTGGCTTGAAG...AGGAATATATCA 1659
Db 1950 AAGATGCTGAAGTGTGAAGATTTTGAAG...GGGATGGCTTA 1991

RESULT 2
US-09-228-986-9
; Sequence 9, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 3239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-9

Query Match 9.2%; Score 178.4; DB 4; Length 3239;
Best Local Similarity 55.3%; Pred. No. 7.4e-35;
Matches 360; Conservative 0; Mismatches 251; Indels 3; Gaps 1;
QY 787 TTCATATACGAGGAGCTAGTACAGCCAAATG...TTCTCCGAGGCAACTTGTATAGGA 846
Db 2026 TTCTCTATGATGAACCTTAAGAGGTGCACCA...AATATTTCTCCGATAGCAATGAATTAGC 2085
QY 847 CAAGCGGGTTCGGTTAGTGTSCAABAAGT...GTGTGCTAGTGGGAAGAGTTGCTGTG 906
Db 2086 TTGGAGGATACGGAAGGTTACAGGGGAGT...TCTTCTGATGCTATATATAGCAATC 2145
QY 907 AAGCAGTTGAAGTTGGAGTGGTTCAGGAG...AGAGGAGTTTTCAGGCGAGGTTGAGATC 966
Db 2146 AAGAGCTCAGCAGGGGTCGATCCAGGTTG...CAACCGAGTTTCAGACAGAAATCGAGCTG 2205
QY 967 ATCAGCAGATTCACCAACAGCATCTGGT...GTCTCTTGTGTTATTCATCGCGGGTCC 1026

Db 2206 CTTTCGGGGTTCATACACAAGAAATCTTTG...TGGCCTCAGGATCTCTTTTCAGCAAGGA 2265
QY 1027 AAAGRTTGTCTGTCTATGAGTTTGTTCCT...TAACAACATCTCAGCTTCACCTCCATGGC 1086
Db 2266 GAGCAGATGTTGGTCTATGATATATGCTTA...ACGGGAGCTCAGGGATAGTTTACAGGA 2325
QY 1087 GAGGAGCGGCTTACAATGGAATGGAGC...CAGCATTTGAAGATTTGCTTTGGATCTGCTAAA 1146
Db 2326 AAATCAGCATTTATCTTGTGAGAGGAGG...CTTCGTATAGCTTAGGTTCGGCTAGA 2385
QY 1147 GGACTTTCCTTATCTTATGAGATTCGAAT...CCTAAATCATTCACCGTGTATCAGGCT 1206
Db 2386 GGACTTTCCTTATCTTATGAGATTCGAAT...CCTAAATCATTCACCGTGTATCAGGCT 2445
QY 1207 TCAACATATTTAGATAGATTTCAAGT...TTGAAGCTAGGTTCGTCTTTGCTCTAAG 1266
Db 2446 ACCAAATCTTGTGGAGCAATCTGACG...GCAAGTCGGGATTTGCGTTTTCCTCAA 2505
QY 1267 ATTCCTTCTGATA---CAAACACGATGTA...TACACAGCTGTGATGGGAACTTTGGGTAC 1323
Db 2506 CTGTATCGGACAGCGGGAAGGGGCACGCT...TTTCGACGCAAGTGAAGGCACTGCGGCTAT 2565
QY 1324 TTGCTCGGGAATAGCTGCAAGCGGAAGCT...CAGGAGAGTCTGACGTTTTCATTT 1383
Db 2566 TTGATCCCAATCTATCATGAGTCAACAG...CTGACAGAAAGAGGATGTGTACAGCTTC 2625
QY 1384 GCGCTTGTGCTTTTGGAGCTCATTAAG...TCTGAGCTGACACCGCTTGA 1427
Db 2626 GGGTGTGCTGCTTGTGAGCTCATCTCA...AGCAACCGATTTGA 2669

RESULT 3
US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7

Query Match 9.0%; Score 174.6; DB 4; Length 2432;
Best Local Similarity 56.5%; Pred. No. 5.7e-34;
Matches 368; Conservative 0; Mismatches 274; Indels 9; Gaps 2;
QY 787 TTCATATACGAGGAGCTAGTACAGCCAAATG...TTCTCCGAGGCAACTTGTATAGGA 846
Db 1070 TTCTCATCAAGAGTACGTTTCCACTAATA...TTTATAGTAGCAAGATATTTTAGGA 1129
QY 847 CAAGCGGGTTCGGTTAGTGTCAACAAGT...GTGTGCTAGTGGGAAAGAGTTGCTGTG 906
Db 1130 GTAGGAGGATATGGAATTTGCTATAAAGG...ATTCCTTACAGATTCGCACTATAGTCAATA 1189
QY 907 AAGCAGTTGAAGTTGGAGT---GGTCAGG...AGAGGAGTTTTCAGGAGAGGTTGAG 963
Db 1190 AAAAGTTGAAGATGGTAATGTGGAGGAG...GAGAAATCAATTTCAACAGAGGTGAA 1249
QY 964 ATCATCAGCAGGTTTCAACAGGATCTGGT...GTCTCTTGTGTTATTCGATCCCGGT 1023
Db 1250 AIGATCAGCTTGTGCTCATAGGAACCTAT...TACGATTTGATGGATTTTGCACACCTCC 1309
QY 1024 GCCAAAGATTCGTTGTCTATGAGTTTCT...TCTTACACAAATCTCG-----AGCTTCA 1077

Db	1310	AGACAGAGCGTCTTGCTCTATCCCTACATGCCAAATGGAAGTGTGGCCTCTGTGCTTAGA	1369
Qy	1078	CTCCATGGCAGGACCGCCCTACAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTGGA	1137
Db	1370	GATCATATTATGGAAGAGCTTGCCCTGGACTGSCCTACTGCRAGCGTATAGCCCTTGA	1429
Qy	1138	TCTCTAAAGGACTTCTTATCTTCATGAGAGATTGCAATCCTAAAAATCATTCACCCGTAT	1197
Db	1430	CGAGCTAGGGGACTGTGTATATTGCAATGAGCAATGTGATCCCAAGATTATCCCGGAT	1489
Qy	1198	ATCAAGGCTTCAACACATATTGATAGATTTCAAGTTTGAAGCTAAGGTTCGTGATTTGGT	1257
Db	1490	GTGAAGCAGCAAAATATATTACTGATGAATATTTTGAAGCTGTGTTGTGAGAGATTTGG	1549
Qy	1258	CTTCTAAGATTGCTTCTGTATACAAACACCATCTATCAACAGCTGTGATGGGAACCTTT	1317
Db	1550	TTACGACAGCTCTTGATACGCGGATTTCTATGTGACTACTGCTGTTCGAGGACCGGTA	1609
Qy	1318	GGGTACTTGGCTCGCGAATAGCTGTCGAAGCGGAAAGCTCAGGAGAAAGTCTGACGTTTC	1377
Db	1610	GGTCACATTGCCCCAGATACCTTTCACGGGACAATCTTCAGAGAAACTGATGATTT	1669
Qy	1378	TCATTTGGCGTTTGGCTTTTGGAGCTCATTTACTGGAGCTGACCGCTTGAT	1428
Db	1670	GGCTTTGGGATTTACTGTGTGGAATCTCATTCAGGACAAAGGCTTTAGAT	1720

RESULTS

```

US-08-265-528-1
; Sequence 1, Application US/08265628
; Patent No. 5821094
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Gorlag, Daphne
; TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE

```

```

? ORIGINAL SOURCE:
? ORGANISM: Brassica napus
? STRAIN: oleifera
? INDIVIDUAL ISOLATE: W1
? POSITION IN GENOME:
? CHROMOSOME/SEGMENT: S-locus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..2574
? PUBLICATION INFORMATION:
? AUTHORS: GÖRLING, DAPHNE
? AUTHORS: ROYSTEIN, STEVEN J.
? TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
? TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
? TITLE: FUNCTIONAL SERINE/THREONINE KINASE
? RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749
? US-08-265-628-1

Query Match      8.68; Score 166.4; DB 1; Length 2749;
Best Local Similarity 56.4%; Pred. No. 6.6e-32;
Matches 352; Conservative 0; Mismatches 266; Indels 5; Gaps 2;

QY 796 GAGGAGCTAGCTAGAGCCACCAATGTTTCTCGAGCGCAACTTGTAGGCAAGGCGGG 855
Db 1558 GAAGCTGTTGTCRAAGCCACCGAAATTTCTCCAAATGTCAACAACTCGGACCAAGTGCGI 1617
QY 856 TTCGGTTACGTGCACAAAGSGTGTTCGCCTAGTGGGAAGAAGTTGCTG-GAAGCAATG 915
Db 1618 TTCGGTATTGTTTACAAGGCTAGATTACTTGTATGTTGGGCAAGAAATTCGGG-AAAAAGGCTA 1677
QY 916 AAGTTGGGAGTGGTCAGGAGAGAGAGGAGTTTTCAGGCAGAGSITTGAGATCATCAGCAGA 975
Db 1678 TCAAAAACGTGCGCTCAAGGAGCTGGTGAGTTTATGTAATGAGSTGAGATTGATCGCGAGG 1737
QY 976 GTTCACCACAGGCACTCGGTGTCCTCTTTGGTTTATTGCATCGCGCGTGCCAAAAGATTG 1035
Db 1738 CTTGACGATATAAACCTTGTCCGAAATCTTTGGCGTTTGCATGTAGGCAGACGAGAAATG 1797
QY 1036 CTTGCTCTATGAGTTTGTTCCTAACACAACTATCGAGCTTCACCTCCATGG- --CGAGGGA 1092
Db 1798 CTGGTATGAGTATTAGAAAATTTAAGCCTCGAATCTTATCTCTCGGAATAAACA 1857
QY 1093 CGGCGTCACATGGAAATGGAGACCCAGATTTGAAGATTGCTCTTGTAATCTGCTAAGSACTT 1152
Db 1858 AGCTCTACGTTAAATTCGAAGACAGATTTCAACATTACCAATGGTGTGGTCTCGAGSACTT 1917
QY 1153 ICTTATCTTCATGAAGATTCGATCCCTAAATCATTCACCGTGATATCAAGGCTTCAAC 1212
Db 1918 TTATATCTTCATCAGACTCAGGTTTAGGATAATCCACAGATATGAAGAATGAAGTAAC 1977
QY 1213 ATATTGATGATTTCAAGTTTGAAGCTTAAGSTTGCTGATTTTGGTCTTTGCTTAAGANTGCT 1272
Db 1978 ATTTTGCTTCATAAAAATATGACACCAAGATCTCGGAATTTGGGATGGCCAGAAATCTTT 2037
QY 1273 TCTGATACACACCGCATGTATCAACAG- --TGTGATGCGAACTTTGGTACTTGGCT 1329
Db 2038 GCAGGGACGAGACTGAAGCTTAACACAAAGAGGTGGTCGGAATCTACGGCTACATGCT 2097
QY 1330 CCGGATACCGTGCACGCGAAAAGCTCCACGAGAAAGTCTCAGCTTTCTTCATTTGGCGTT 1389
Db 2098 CCGGAGTACCGAATGGATGGGGTATCTCGGAAAATACATGTTTTCAGTTTGGAGTC 2157
QY 1390 GTGCTTTGGAGCTCAATTACITGGA 1413
Db 2158 ATTGTTCTTGAATTTGTAGTGA 2181

```

DECEMBER 5

RESULTS 3
US-08-881-706-1
; Sequence 1, Application US/08881706
; Patent No. 6245969
; GENERAL INFORMATION:
; APPLICANT: Chong, Joame

```
; APPLICANT: Li, Jianming
; TITLE OF INVENTION: Receptor Kinase BIN1
; FILE REFERENCE: 07251/022001
; CURRENT APPLICATION NUMBER: US/08/881,706
; CURRENT FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4104
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)...(3687)
US-08-881-706-1

Query Match      8.4%; Score 162.4; DB 4; Length 4104;
Best Local Similarity 55.4%; Pred. No. 7.9e-31;
Matches 382; Conservative 0; Mismatches 296; Indels 12; Gaps 3;

QY 788 TCACATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCCGAGGCGAATCTGTTAGGAC 847
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 2708 TCACGTTCCGATCTCTCTAGGCTACCAATGGTTTCCATATGATGATCTGTTGGTT 2767
QY 848 AAGCGGGTTCGTTACGTGCGACAAAGGTGTGTGCTGCTAGTGGGAAAGAGTTGCTGTGA 907
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 2768 CTGGTGGTTTGGAGATGTTTACAAAGCGATTTGAAAGATGGAAGCGGGTGGCTATCA 2827
QY 908 AGCAGTTTAAAGTTGGAGTGTCTAGGAGAGGAGGTTTTCAGGAGAGTTGATGATCA 967
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 2828 AGAAATGATTCATGTTAGCGGTCAAGGTGATAGAGATTCATGCGGAGATGGAAACCA 2887
QY 968 TCACAGAGATTCACACAGCAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1027
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 2888 TTGGGAAGATCAAAATCAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2947
QY 1028 AAGATGTTCTGCTATGATGTTTCTTCAACAAATCTCGAGCTTCACCTCCATGGC- 1086
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 2948 AGCGGCTCTCTGTTAATGAGTTATGAAATGGAAGTTAGAGATGTTTGCAGAGACC 3007
QY 1087 -----GAGGAGCGGCTCAATGGAATGAGGAGCAGCAGATGGAATGCTCTTGGATCTG 1141
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3008 CCAAGAAAGGTGGGTGAAACTTAAATTTGTCACACGCGGGAAGATTCGATAGGATCAG 3067
QY 1142 CTAAGAGGACTTCTTATCTCTATGAGATGCAATCTTAAATCAATCAATCAATCAATCA 1201
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3068 CTAGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3127
QY 1202 AGGCTTCAACATATGATAGATTTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTTG 1261
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3128 AATCCAGTAAATGTTGTTGATGAGAAATTTGGAAGCTCGGTTTCAGATTTTGGCATGG 3187
QY 1262 CTAAGATGCTTCTGATCAACACGATGATACACAGT---GTGATGGGAACCTTTG 1318
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3188 CGAGGCTGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3247
QY 1319 GGTATGCTGCTCGGAATAGCTGCAAGCGGGAAGCTACGAGAGATCTGAGCTTTTCT 1378
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3248 GTTAGTCTCTCCAGATGATACCAAGTTTCAAGTTTCAAGATGATGATGATGATGATGAT 3307
QY 1379 CATTTGGGCTTGTTGTTTGGAGCTCATTTAGTGGAGCTGAGACCGCTTGTATGCCAATG 1438
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3308 GTTAGGCTGTGCTCTACTCGAGTACTCAGCGGTAACGCGGATTCACCGGATTCACCGGAT 3366
QY 1439 TCTATGTAGATGACAGCTTAGTTGACTGGG 1468
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 3367 --TTTGGAGATAACAACTTGTGGATGGG 3394

RESULT 6
US-07-717-331F-9
; Sequence 9, Application US/0717331F
; Patent No. 5484905
```



```

QY 1277 ATACAAACGCGCATGTATCAACACGCTGTGTGATGGAACTTTGGTACTTGGTCTCCCGAAT 1336
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2042 ACGAACGGAAGTACACAAATGAAGTGTGCGAACATACGCTACATGTCCCGGAAT 2101
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1337 ACCTGCAAGCGGAAGCTACGGAAGTCTGACGTTTCTCATTTGGGCTTGTGCTTT 1396
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 ACCTCAATGTATGGATATCTCGGAAATACAGATGTTTCAGTTTGGAGTCATGTC 2161
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1397 TGGAGCTCATTACTGGA 1413
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2162 TTGAATGTGTAGTGA 2178
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-07-717-331F-1
; Sequence 1, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ. ID. NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-07-717-331F-1

Query Match: 7.5%; Score 146.6; DB 1; Length 2833;
Best Local Similarity 54.6%; Pred. No. 5.6e-27;
Matches 337; Conservative 0; Mismatches 274; Indels 6; Gaps 2;

QY 803 TAGCTAGACCCCAATGGTTTCTCGAGCGCACTGTAGGACAAAGCGGTTGCGTT 862
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1562 TTGTAAAGCCCGGAAATTTCTCCAGCTGTAAACAACCTCGCAACAGTGTGTTTGGTA 1621
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 863 ACCTGCACAAGGISTGTTGCCFAGTGGGAAAGTGTCTGTGAAGCACTGTAAGTTG 922
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1622 TTGTTTACAGGGAAGATTACTTACGCGGAAGAAATTCAGTAAAGGCTATCAAGA 1681
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 923 GGAGTGGTCAGGAGAGAGGAGTTTCAGGACGAGGTTGAGATCAATCAGCAGAGTTCACC 982
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1682 CGTCAGTTCAGGAGACTGAGTATGATGAGGTTGACACTAATTCGAGGCTTCAGC 1741
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 983 ACAGGATCTGGTGTCTCTGTGGTTATGTATGATCCCGGTCGCAAGATTCCTGTCT 1042
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1742 ATATAAACCTTGTTCAGG-TCTTGGCTGTGTGATGTAAGGAGATGAGAAGTGTGATAT 1801
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1043 ATGASITGTTCCTAACAAACAAATCTCGAGCTTCACCTCCATGGCGA---GGACGGCCTA 1099
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1802 ATGAGTATTTGGAATTTAAAGCCTTGATTTCTATCTTTTGTAAACCCGAGGTCTA 1861
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1100 CAATGGAAATGAGCAGACAGATTGAGATTGCTCTTGGATCTGCTAAAGGACATTTCTTATC 1159
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1862 AGCTAATTTGAANTGAGAGATTCGACATTTACCAATGGTGTCTCGAGGCTTTTATATC 1921
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1150 TTCATGAAGATTGCAATCCTAAATCATTCACCGTGATACAGGCTTCAACACATATGA 1219
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1922 TTCATCAAGACTCAGGTTTAGGATAATCCAGAGATTGGAATTAAGTAACATTTGTC 1981
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1220 TAGATTTCAGCTTTCARCTAAGCTGTCTGATTTGTTGTTCTCTAAGA---TTGCTTCTG 1276
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1982 TTGACAAAATATGATCCCAAGATCTCGGATTTTGGATGGCCAGGATATTTGAAAGG 2041
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1277 ATACAAACACGCATGTATCAACACGCTGTGATGGGAACCTTTGGTACTTGGCTCGGAAT 1336
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2042 ACGAACGGAAGCTAACACAATGAAGTGTGTCGGAACATACGGCTACATGTCCCGGAAT 2101
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1337 ACCTGCAAGCGGAAGCTACGGAAGTCTGAGGAAAGTCTGACGTTTCTCATTTGGGCTTGTGCTTT 1396
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 ACCTCAATGTATGGATATTTCTCGGAAATACAGATGTTTTCAGTTTGGAGTCATGTC 2161
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1397 TGGAGCTCATTACTGGA 1413
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2162 TTGAATGTGTAGTGA 2178
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-08-587-680A-24
; Sequence 24, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774

```


RESULT 10

US-09-228-986-1

; Sequence 1, Application US/09228986

; Patent No. 6359198

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Niels

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000/1020

; CURRENT APPLICATION NUMBER: US/09/228,986

; CURRENT FILING DATE: 1999-01-12

; NUMBER OF SEQ ID NOS: 130

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 2389

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-228-986-1

Query Match

Best Local Similarity 50.9%; Score 138.2; DB 4; Length 2389;

Matches 425; Conservative 0; Mismatches 383; Indels 27; Gaps 3;

QY	831	GGCGAAGTTGTTAGACAGCGGGTTCGGTTACGTCACAAAGGTGTTGCCCTAGTGG	890
DB	1189	GGCTGAATGCTAGGAAGGGTAACCTTGGAAAGCCCTTATAAGGCTATATGGAAGATGG	1248
QY	891	GAAGAAGTTGCTGTGAAGCAAGTTGAAAGTTGGGAGTGCTCAGGAGAGAGGGAGTTTCA	950
DB	1249	ATCTGTTGCTGTCAGAGAGTTGAAGAGTTTGTATGTTATGGAAGAGAGAGTTTCA	1308
QY	951	GGCAGAGTTGAGATCATCAGCAGAGTTTCACACAGGAGTCTGTGTCCTTTGGTTA	1010
DB	1309	GCACATAAGGAGTCTATGGAAGTTTGAAGATCAAAATGATTAATTTAAGGSCCTA	1368
QY	1011	TTCATCCCGGTGCCAAAGATTCCTGTTCTATGAGTTTGTCTTCAACAAATCTCA	1070
DB	1369	TTATTTCTAGAGATGAAGAGCTTCTGCTATGACTATATGCCAAATGGAGTCTGTA	1428
QY	1071	CTTTCACCTCCATGCGAGGAGCGCCCT-----ACAATGGAATGGAGCACCGAAT	1121
DB	1429	TGCACCTTCTCAIGTACCGAGGACCTGGTAGACACCCCTAGACTGCACACTAGAT	1488
QY	1122	GAAATGCTCTTGGATCTGTAAGGACTTCTTATCTCTATGAGTTCATGAGATTCGA	1178
DB	1489	GAGATAGCACTTGGTGCAGCTAAAGGCTTGGCAATTTATCACAGCCACTGCAAGTCACC	1548
QY	1179	TAAATCAATCCCGTGTATCAAGGCTTCAACATATGATTTCAAGTTTGAAGC	1238
DB	1549	CAAGTTGGCCACGGAAACATCAAGTCTTCCAAATTTTGTATGAGAGATGGGAATGC	1608
QY	1239	TAAAGTTGCTGATTTGCTTGTCTAAGATTCCTTGTATACAAACAGCCATGTATCAAC	1298
DB	1609	TTCATTTACAGACTTTGGCTTSCACTGCTG-----TGAGTCCATCACT	1653
QY	1299	ACGTGATGGAGACCTTTGGGTACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCAC	1358
DB	1654	TGCTGCTTACAGAAATGTTGGGCTACACTGCTCCTGAGCAAGCTGCAACAAAGAGATATC	1713
QY	1359	GGAGAGTCTGACCTTTCTCATTTGGCTTGTCTTTTGGAGTCTATGAGTGGAGCTG	1418
DB	1714	CCGAAAGCGGACGTGTACAGCTTTGGAGTTTACTGCTTGAATGCTAACTGGCAAGGC	1773
QY	1419	ACCGTTGATGCCAACATGCTATGTAGATGACAGCTTACTGTGCTGGCAGCAGCAAT	1478
DB	1774	CCCTGTTCAAGCTCACATGCAAGAGGATTCATTCAGCTATAGTCTTCCAGGTGGGT	1833
QY	1479	CTTTACCGGAGCATCTGAGCAAGAGAGCTTTGAGGTTTACTGATGCAAGATGAATAA	1538
DB	1834	GCAGTCCATTTTCCAGAAAGATGACATCAGAGTATTTGATATATGATGATGAGATT	1893

QY	1539	TGSGTATGACAGAGAGAGATGGCTCCATGTTGCTGCTGGCGGTGTTGTTCCGCA	1598
DB	1894	CAAAAACATTGAAGAGGAACTAGTAGCATGTTTGTCTTGTCTGCTCACA	1953
QY	1599	TCAGCTCGCCGACAGACTCGCATGAGCAGATGTGCGTGGTTAGAAGAAAT	1653
DB	1954	ATCTCTCAGCAACGCCCTAAATGAGCCATGTTGAGGTTGATTCAGACATT	2008

RESULT 11

US-C9-228-986-10

; Sequence 10, Application US/09228986

; Patent No. 6359198

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Niels

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000/1020

; CURRENT APPLICATION NUMBER: US/09/228,986

; CURRENT FILING DATE: 1999-01-12

; NUMBER OF SEQ ID NOS: 130

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 2336

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-228-986-10

Query Match

Best Local Similarity 50.4%; Score 130.8; DB 4; Length 2336;

Matches 346; Conservative 0; Mismatches 33; Indels 3; Gaps 1;

QY	787	TTACATACAGCAGAGCTAGCTAGAGCCCAATAGTTTCTCCGAGGCGAACTTTTACGA	846
DB	1024	TTGATTACAGAGAGCTGCTATCCACCGCGAGGTTTCGCGACGAAACCTTTTGGG	1383
QY	847	CAAGCGGGTTCGTTACGTGCAACAAAGGTGTTCCTAGT---GGGAAAGAGTTGCT	903
DB	1084	TACGGGGGATTTGGCATGTTTACAAGGGTTTCTCCCCAGGAGCGGCCAAGAAGTCGA	1143
QY	904	GTGAACAGTTGAAAGTTGGAGTGTTCAGGGAGAGGAGGAGTTTCAGCGAGAGTTGAG	963
DB	1144	GTGAATGTATACAGCGAGTTCAAGGAGGATTAAGGGTTTGTTCAGAGATCTCA	1203
QY	964	ATCATACAGAGATTTACACAGGCTATCTGTTCTCTTTTGGTTTATTCATCGCGGT	1023
DB	1204	AGCATGGGCGGTACAGCACCGAACCTGGTTCAACTCCGAGGATGGTCCGAGGCAAT	1263
QY	1024	GCCAAAGATTCCTTGTCTATGAGTTTGTCTTAAACAAATCTCAGCTTCACTCCAT	1383
DB	1264	ACACAGCTATTCATCGTTTACGACTACATGCCAACGGAAGCCTTCAACTCATCTTC	1323
QY	1084	GCGAGGGAGCGCTTACATGGAATGGACACAGATTTGAAGATTCCTCTTGGATCTGTT	1143
DB	1324	GGTAGTCCGCAACAGCTCTCTCGGTGCTATCGCGATACGCGATCTCAAGGCGTAGCA	1383
QY	1144	AAAGGACTTCTTATCTTATCATGAAGATTCGAATCTTAAATCATTCACCGTGTATCAAG	1203
DB	1384	GCGGGGCTGCTATCTCTGACAGGCAATGGGAGAGAGGGTCTGCCACAGGACATTAAG	1443
QY	1204	GCTTCAACATATTCATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTGCT	1263
DB	1444	TCGAGCAACGCTGCTTGTGATTCGAGTTTCAACGGGCGCTTAGTACATTCGCGCTCGCT	1503
QY	1264	AGATTTGCTCTGTATACAAACAGCATGTATCAACAGCTGTGATGGAGACCTTTGGGTAC	1323
DB	1504	CGGCTGTATGATCAGTGAGATTCGAGAGACATATGTGTGAGGACGTTGGGGTAC	1563
QY	1324	TTGGCTCCGGAATACGTCGAAGCGGAAAGCTCACGAGAGTCTGACGTTTCTCATTT	1383
DB	1564	ATAGCAGCGAGTTTGATACAAACGSGGAGGAGGCACTCTCTAGCTCGGACGCTTCAGCTTC	1623


```

RESULT 14
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nicuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA

```

```

1  RESULT 15
2  US-08-781-891-208
3  ; Sequence 208, Application US/08781891
4  ; Patent No. 6090520
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Fu, Ying-Hui
7  ; APPLICANT: Yu, Chang-En
8  ; APPLICANT: Oshima, Junko
9  ; APPLICANT: Mulligan, John T.
10 ; APPLICANT: Schellenberg, Gerald D.
11 ; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
12 ; TITLE OF INVENTION: WERNER'S SYNDROME
13 ; NUMBER OF SEQUENCES: 209
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: SEED AND BERRY LLP
16 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
17 ; CITY: Seattle
18 ;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 07:27:14 ; Search time 5106 Seconds
(without alignments)
11080.269 Million cell updates/sec

Title: US-10-086-464-1
Perfect score: 1944
Sequence: 1 atgtctctggcgccgtctcc.....atagtgacctctcttttaa 1944

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rtd.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. Nc. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1944	100.0	1944	6	AX088876	AX088876 Sequence
2	1944	100.0	2189	8	AY028699	AY028699 Brassica
3	1369.4	70.4	2098	8	AY128792	AY128792 Arabidops
4	1369.4	70.4	2116	8	AY056788	AY056788 Arabidops
5	1369.4	70.4	2188	8	AY059901	AY059901 Arabidops
6	1369.4	70.4	2190	8	AY093065	AY093065 Arabidops
7	1367.8	70.4	2257	8	AF370509	AF370509 Arabidops
8	1341.2	69.0	2324	8	AY089024	AY089024 Arabidops
9	516.2	26.6	1902	6	AX088882	AX088882 Sequence
10	516.2	26.6	1939	6	AX088881	AX088881 Sequence
11	46.6	23.0	79706	8	AB020746	AB020746 Arabidops
12	427.2	22.0	2025	6	AX088885	AX088885 Sequence
13	427.2	22.0	2104	6	AX088884	AX088884 Sequence
14	413.6	21.3	2196	6	AX088888	AX088888 Sequence
15	413.6	21.3	2261	6	AX088887	AX088887 Sequence
16	405.8	20.9	2046	8	AY113039	AY113039 Arabidops
17	405.8	20.9	2616	8	AF424623	AF424623 Arabidops
18	405.8	20.9	2735	8	AY075681	AY075681 Arabidops
19	402.8	20.7	1615	8	AY060577	AY060577 Arabidops
20	399.6	20.6	2158	8	AY113877	AY113877 Arabidops
21	399.6	20.6	3064	8	AY035076	AY035076 Arabidops
22	353.2	18.2	174752	8	AC069324	AC069324 Oryza sat
23	305.6	15.7	2196	6	AX088879	AX088879 Sequence
24	305.6	15.7	2231	6	AX088878	AX088878 Sequence
25	301.8	15.5	53393	8	AF000382	AF000382 Arabidops
26	240.8	12.4	2062	8	AY096404	AY096404 Arabidops
27	240.8	12.4	2169	8	AY063963	AY063963 Arabidops
28	229.6	11.8	75508	8	AB026654	AB026654 Arabidops
29	228.6	11.8	2837	8	AY056259	AY056259 Arabidops
30	216	11.1	145120	8	AP003434	AP003434 Oryza sat
31	215.8	11.1	2414	8	AF106957	AF106957 Nicotiana
32	211.5	10.9	2266	8	AY133858	AY133858 Arabidops
33	211.6	10.9	2936	8	AY091071	AY091071 Arabidops
34	207.4	10.7	3674	8	AY094451	AY094451 Arabidops
35	204.6	10.5	1419	8	AF332429	AF332429 Arabidops
36	204.6	10.5	1440	8	AY093964	AY093964 Arabidops
37	204.6	10.5	1754	8	AF367265	AF367265 Arabidops
38	196.6	10.1	1158	8	AY064968	AY064968 Arabidops
39	195.4	10.1	2077	8	ZMA277703	ZMA277703 Zea mays
40	194.8	10.0	2004	6	AX461250	AX461250 Sequence
41	194.8	10.0	113566	8	AC002330	AC002330 Arabidops
42	194.8	10.0	194892	8	ATVCHRIV6	ATVCHRIV6 Arabidops
43	194.6	10.0	1750	8	AY087491	AY087491 Arabidops
44	192.4	9.9	2625	8	AY125515	AY125515 Arabidops
45	192.2	9.9	114367	8	ATF10ML0	ATF10ML0 Arabidops

ALIGNMENTS

RESULT 1
AX088876
LOCUS AX088876 1944 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 1 from Patent WO0114563.
ACCESSION AX088876
VERSION AX088876.1 GI:13397639
KEYWORDS
SOURCE rape.
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1944)
AUTHORS Goring,D. and Silva,N.
TITLE Proline-rich extensin-like receptor kinases

Db	734	AAATCATGTAGTCAGCTCATTGCGCACACCTAAGCCCTCCATCTCCACACGAAACCTCTCT	793
QY	676	-----CCACCTCCACCCGCGCTTTTCATGAGCAGCAGCGCGCGCTCCGACTACTCGGACGCT	732
Db	794	CGGCCACCTCCACCCAGCATTTCATGAGTACGAGTGGTGTCTGACTATTCGGATCTT	853
QY	733	CCAGTCTCTCCACCGCTCCAGGCTTGTGTTAGGCTTCTCCAAAAGCAGCTTCACA	792
Db	854	CGGTTCTTCTCCACCACTCCAGGCTTGTGTTAGGCTTCTTCTAAAAGCAGCTTCACT	913
QY	793	TACGAGGAGTACCTAGAGCCACCAATGTTCTCGAGCGCAACTTGTAGCAGAGCC	852
Db	914	TATGAGGATTCTCAGAGCTACTATGAGCTTCTCTGAGGCTAATTTGTAGCAGAGA	973
QY	853	GGGTTGGTTACGTGCACAAAGGTGTGTTGCTAGTGGGAAAGAGTGTCTGTGAAGCAG	912
Db	974	GGGTTGGTTATGTCATAAAGGTATATGCTAGTGGGAAAGAGTGTCTGTGAAGCAG	1033
QY	913	TTGAAGTTGGAGTGGTTCAGGAGAGAGGGAGTTTCAGGAGAGTTGAGATCATCAGC	972
Db	1034	TTGAAGCTGGTAGTGGTTCAGGAGAGAGAGTTTCAGGAGAGTTGAGATCATCAGC	1093
QY	973	AGAGTTCACACAGCAGCTCTGCTCTCTGTTGTTATTTSCATCCCGGTGCCAAAGA	1032
Db	1094	AGAGTTCATCAGAGCATTGGTTCTCTATTTGTTATTTGATGGCGGTGTCTCAAGA	1153
QY	1033	TGCTTGTCTATGAGTTGTTCTCTTAACAACAATCTGAGCTTCACCTCCATGGCGAGGA	1092
Db	1154	TTACTTGTCTATGAGTTGTTCTCAACAACAATCTGAGCTTCACCTCCATGGTAAAGGA	1213
QY	1093	CGGCTTACAATGGAATGAGCAGCAGTGTGAGTGTCTCTGTTGATCTCTTAAGGACTT	1152
Db	1214	CGGCTTACATGGAATGAGTGTGAGTGTGAGTGTCTCTGTTGATCTCTTAAGGACTT	1273
QY	1153	TCTTATCTTCATGAAGATTGCAATCTTAAATCAATTCACCGTGATATCAAGCTTCAAC	1212
Db	1274	TCATATCTTCATGAAGATTGCAATCGGAAATCAATTCACCGTGATATCAAGCTTCAAC	1333
QY	1213	ATATGTATGATTTCAAGTTGAAGTGAAGTGTGCTGATTTGGTGTCTTAAGATTGCT	1272
Db	1334	ATATGTATGATTTCAAGTTGAAGTGAAGTGTGCTGATTTGGTGTCTTAAGATTGCT	1393
QY	1273	TCTGATACAAACACGATCATCAACAGCTGTGATGGGAGGCTTGGTGTCTTAAGATTGCT	1332
Db	1394	TCTGATACAAACACGATCATCAACAGCTGTGATGGGAGGCTTGGTGTCTTAAGATTGCT	1453
QY	1333	GAATACGCTGCAAGCGGAAAGCTCACGGAGAGTCTGAGCTTCTCATTTGCGGTGTG	1392
Db	1454	GAATACGCTGCAAGCGGAAAGCTCACGGAGAGTCTGAGCTTCTCATTTGCGGTGTG	1513
QY	1393	CTTTTGGAGCTTACTTGGAGCGTGGACCCCTTGATGCCAACAATGCTATGATGATGAC	1452
Db	1514	CTTTTGGAGCTTACTTGGAGCGTGGACCCCTTGATGCCAACAATGCTATGATGATGAC	1573
QY	1453	AGCTTAGTTGCTGGGACGACCATTTGCTTACCGAGCATCTGACAGGAGCATTTGAG	1512
Db	1574	AGCTTAGTTGCTGGGACGACCATTTGCTTACCGAGCATCTGACAGGAGCATTTGAG	1633
QY	1513	GGTTAGCTGATCAAGATGAATAATGGGTATGACAGAGGAGATGGCTCGCATGTT	1572
Db	1634	GGTTAGCTGATCAAGATGAATAATGGGTATGACAGAGGAGATGGCTCGCATGTT	1693
QY	1573	GCTTGTCTGGGCTTGTGTTGCGCATTCAGCTCCCGCAGCAGCTTCGATGAGCAGATT	1632
Db	1694	GCTTGTCTGGGCTTGTGTTGCGCATTCAGCTCCCGCAGCAGCTTCGATGAGCAGATT	1753
QY	1633	GTCGCTGCTGTAGAGGAATGATCATCTGTCAGATCTTACGAGGAGATGAGACCGT	1692
Db	1754	GTACGGGCTGTAGAGGAATGATGATGCTGTCTGTCATCTTACGAGGAGATGAGACCGG	1813
QY	1693	CAAGCAATGTATACGCTCATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1752
Db	1814	CACAGCAAGGTATACGCTCATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1873
QY	1753	GAAGACATGAAGAGTTTAGGAAATGCGCTTGGAACTCAAGACTCAAGACGCCACGGGT	1812
Db	1874	GAGACATGAAGAGTTTAGGAAATGCGCTTGGAACTCAAGACTCAAGACGCCACGGGT	1933
QY	1813	GAGTACAGTAAATCCGACAGTACTGAGTGTACCCGCTCTGTTTCAAGCAGCAGGAGG	1872
Db	1934	GAGTACAGTAAATCCGACAGTACTGAGTGTACCCGCTCTGTTTCAAGCAGTGAAGGT	1993
QY	1873	CAACACACAGCGGAATGAGATGGGGAAGATTAAGAACCCGCTCAAGTTATAGTGA	1932
Db	1994	CAACACACAGCGGAATGAGATGGGGAAGATTAAGAACCCGCTCAAGTTATAGTGA	2053
QY	1933	CTTCTCTTTAA 1944	
Db	2054	CCCTCTTTAA 2065	
RESULT 8			
LOCUS	AY089024	2324 bp mRNA linear	PLN 21-JUN-2002
DEFINITION	Arabidopsis thaliana clone J7909 mRNA, complete sequence.		
ACCESSION	AY089024		
VERSION	AY089024.1	GI:21407798	
KEYWORDS	FLI_CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
TITLE	1 (bases 1 to 2324)		
JOURNAL	Haas,B.J., Volkovsky,N., Town,C.D., Troupkan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,C. and Salzberg,S.L.		
REFERENCE	Full-length messenger RNA sequences greatly improve genome annotation		
AUTHORS	Genome Biol. (2002) In press		
TITLE	2 (bases 1 to 2324)		
JOURNAL	Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
REFERENCE	Full-length cDNA from Arabidopsis thaliana		
AUTHORS	Unpublished		
TITLE	3 (bases 1 to 2324)		
JOURNAL	Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA		
TITLE	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.		
COMMENT	Location/Qualifiers		
FEATURES	1..2324		
SOURCE	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/clone="17909"		
BASE COUNT	583 a 566 c 483 g 692 t		
ORIGIN			
Query Match:	69.0%; Score 1341.2; DB 8; Length 2324;		
Best Local Similarity	82.8%; Pred. No. 8.3e-260;		

Rosidae; eursids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1902)

Goring,D. and Silva,N.
Proline-rich extensin-like receptor kinases.
Patent: WO 0114563-A 7 01-MAR-2001;
Goring, Daphne (CA) ; Silva, Nancy (CA)
location/Qualifiers
1..1902
/organism="Arabidopsis thaliana"
/db xref="taxon:3702"
1..1902
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34705.1"
/db_xref="GI:13397646"

/translation="MASSPESEAPTNTSSPPSPNSTTSSPPSPNSTTSSPPAPSDPSPTPEQD
SSSPPDSTSPAPQAPPNPNSNDSPSPSGSGGERGNNGNNTTALIVGLVGA
SPPSRNEDNGRSRPPGTGSRSDPSPSPSGSGGGGRGNNTTALIVGLVGA
GLIMVIIICLRKKRRKDFYEPKMKNQYYIYNNNNNASQNYPNHLSQQQN
QQTGGGGGGGGPRRPDRMTSGEDSSMYSGSRPLVLPSPPALAGFNKSLFTTQE
LAAATGFDANLLGGFGFYHKVPGSKWEAKSLAGSGRGREFCAFQAVDIISR
VHRUYVLVIWQCIADQGMVLYEVPNKLTLEVHLHGKNIPLVMEFSTRCLRIAAGAAG
LAYLEDCHPRIITHRIKSNAILIDNFDMAYADFSLKITSDNTHVSTRVMGTGY
LAFNEASKLTKESDYISGVMLLEITGRPVDSNTMDDTLVDMFLMARLEALD
GFNEILADRLKGINIPQEMARMWTCAAASIRHSKRKPMSQIVRALGEVEISDALN
EYKPGHNSYVIGLASSYSQTSYNADMKKFERQIALISSQEFVPSDECTSSNSRDMD
GTRKSPTPPK"

BASE COUNT 523 a 457 c 439 g 483 t

ORIGIN

Query Match 26.6%; Score 516.2; DB 6; Length 1902;
Best local Similarity 67.1%; Pred. No. 1.le-93;
Matches 747; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

QY 730 CGTCCAGTCTTCCTCCACCGCCTCCAGGCGTTGTGTAGCTTCTCCAAAAGCACCATTTC 789
Db 757 CGCCCCAGTTTACTCTCTCCCTTCGCCCTGTCTTAGCCCTCGGANTCAACAAGACACTTTT 816
QY 790 ACATACAGAGAGCTAGCTAGACGCACCAATGGTTTTCTCGAGGCGGAAGTTGTTAGACAA 849
Db 817 ACTTACCAGAGAGCTTGGCGTGCAACAGGAGGGTTTTACGGATGCTTAACCTTTGGGACAG 876
QY 850 GGCGGGTTGGTTACGTGCACAAAGTGTGTTCCCTAGTGGGAAGAAGTTCGTCGTCAAG 909
Db 877 GGAGGATTTGGGTATGTCATAAAGAGTCTTCCTTAGCGGGAAGAAGTAGCAGTTAAG 936
QY 910 CAGTTCAAGATTGGGAGTGGTACAGGACAGAGAGGAGTTTCAGGCAGAGTTGAGATCATC 969
Db 937 AGTTTAAAAAGGGTACGGACAGAGAGAGAGGAGGAGTTTCAAGCTGAGTTCGATATCAIT 996
QY 970 AGCAGAGTTTCCACACAGGAGCATCGGTGTCTCTGTGGTATTATTCATCGCGGGGCCAAA 1029
Db 997 AGCGGTGTCATCATCGGTATCTTGTTCTTTGGTGGATATTGCATAGCTGAAGACAG 1056
QY 1030 AGATTGCTTGTATGAGTTTGTTCCTAACAAACATCTCGAGCTTCAACCTCCATCGCGAG 1089
Db 1057 AGGATGTGTTTATGAGTTTGTTCCTAACAAAACATTGGGAATATCATCTTCATGGAAA 1116
QY 1090 GGACGGCCCFACATGNAATGGACACACAGATTGAAGATGCTCTTGGATCTGCTAAAGGA 1149
Db 1117 AATCTTCCGGTAATGAGGTTCTCCAATGGTTCCGTCATCGCCTAGGTGTCGGAAGGA 1176
QY 1150 CTCTCTATCTTCATGAAGATTCGAATCTCAATTAATCAITCAACCGTGATATCAAGGCTCA 1209
Db 1177 CTCGCTTACCTTCACGAGAGCTGCCATCTCCGGATCATTCACCGGACATCAAGTCTGCA 1236
QY 1210 AACCATATGATAGATTCAAGTTTGAAGCTTAAGCTTGCCTGATTTTGGCTTGTCAAGATT 1269
Db 1237 AATATCTCTTGACATCAACTTTTGATGCTATGCTGCTGCTGATTTTGGATTAGTATTA 1296
QY 1270 GCTTCTGATACAAACAGCATGATATCAACACGTGTGATGGAACCTTTGGGTACTTGGCT 1329
Db 1297 ACATCTGATACAAACACTCATGATATCTACTCGTGTGATGGAACCTTTGGATATCACTGCT 1356

```
QY 850 GCGGGTTCGGTTACGTGACAAAGGTGTGTGCTAGTGGGAAAGAGTTCGTGAAG 909
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 910 CAGTTGAAAGTTGGAGTGGTTCAGGAGAGAGGGAGTTTCAGCAGAGAGTTCAGATCATC 969
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 970 ACAGAGTTCACACAGGATCGTGTGTCCTTTGTTGGTATATGATCCCGGTGCAAA 1029
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1014 AGCCGTGTCATCATCGGTATCTTCTTTGGTGTATGATGCTGATGACAG 1073
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1030 AGATGCTGTCTATGATGTTTCTCTAACAAATCTGAGCTTCACTCATGCGGAG 1089
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1074 AGGATGTTGGTTATGAGTTTCTCTAACAAATCTGAGCTTCACTCATGCGGAG 1133
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1090 GGAGCGCTTACATGGAATGAGAGGACCAAGATTAAGATTCCTGCTGCTAAAGGA 1149
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1134 AATCTTCGGTAAAGAGTTCCTCACTAGGTTCGGTATGCTGCTAGGTGCTGCGAAGGA 1193
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1150 CTTCCTTATCTTCAAGAGATTCATCTTAAATCAATTCACCGGTGATATCAAGGCTTCA 1209
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1194 CTCGCTTACCTTCAGAGAGACTGCCATCTCGGATCATTCACCGGACATCAAGTCTGCA 1253
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1210 AACATATTGATAGATTTCAAGTTTGAAGCTAAGGTTCCTGATTTTGGTCTTCTAGATTT 1269
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1254 AATATCTTCTGCACTTCACTTGTATGCTATGTTGGCTGATTTTGGATTAAGTATTA 1313
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1270 GCTTCGTATACAAACAGCATGATATCAACAGCTGTATGGGAAACCTTTGGTACTTGGCT 1329
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1314 ACATCTGATACAAACACTATATCTACTCTGTGTGATGGGAACCTTTCGATATCTAGCT 1373
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1330 CGGGAATAGCTGCAAGCGGAAGCTCACGAGAGCTGACGTTTCTCATTTGGCTTT 1389
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1374 CAGAAATAGCTTCAAGCGGTAAATACCGAGAAATCGATGTTTCTTACCGAGTT 1433
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1390 GTGCTTTTGGAGCTCATTAAGCTGAGCTGAGCGGTGATGACCAACAAATGCTATGATGAT 1449
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1434 ATGTTATTGGAACITATACTGAAAGAGCGGTTGAT--AATAGCATCACCATTGAC 1490
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1450 GACAGCTTAGTGTAGTGGGAGCAGCATGCTTAAACCGAGCATCTGAGCAAGAGACTTT 1509
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1491 GACACCTTAGTGTAGTGGGCTGCGCTCTTATGCTCGCGCTAGAGATGGAACCTTT 1550
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1510 GAGGCTTAGTGTAGTGAAGATGAATATGTTGTTATGACAGAGAGAGATGCTGCGAG 1569
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1551 AATGAGCTGCGAGATGCGAGGCTTGAAGCAACTACACCCGCAAGAAATGCTCGAATG 1610
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1570 GTTGTCTGTGCTGGGCTTGTGTCGCCATTCAGCTCGCGGAGACCTCGCATGAGCCAG 1629
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1611 GTGACTTGTGCGCTGTAGCATTCGTATTCGGGGGTAAACGTCCTCAAGATGAGCCAG 1670
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1630 ATTGTGCTGCTGTAGAGAGAAATGATACATGCTCAGATCTTAAGCAAGGAGAGACCA 1689
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1671 ATAGTAGAGCGTTAGAGAGAGAGAGTTCCTTAGATGCTTAAACGAGAGGTTGAAGCCA 1730
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1690 GGTCAAGCAATGATACAGCTCATACGAGAGAGAGAGAGAGAGATGCTGAGCCAGTAC 1749
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1731 GGACACAGTACGTTTACGGTTCATTTGGAGCAAGCTCGGATTTATGACACATCTTAC 1790
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1750 AATGAAGACATGAAGAACTTTAGGAAATGACCTTTGGAATCAAGAGATCAACCCACG 1809
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1791 AATGACAGATGAAGAAATTCAGACAGATAGCTTTGTCGAGCCAGAAATTCACATCTAGT 1850
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1810 GGTGAGTACATATCCGACAGTACATATGGA 1842
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1851 GACTGTGAAGGAACATCTAG--AATGATCTAGA 1883
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

AB020746

LOCUS

AB020746 79706 bp DNA linear PIN 27-DEC-2000

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MOB24.
AB020746 BAB00014
AB020746.1 GI:3985949

ORGANISM

Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
clone:MOB24.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
AUTHORS
TITLE

1 (sites)
Kaneko,T., Kato,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,895 bp regions covered by 90 P1, TAC
and BAC clones
DNA Res. 7 (3), 217-221 (2000)

JOURNAL
MEDLINE
REFERENCE

2 (bases 1 to 79706)
Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
Direct Submission
Submitted (03-DEC-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakam@kazusa.or.jp,
tel:81-438-52-3935, fax:81-438-52-3934)

AUTHORS
TITLE
JOURNAL

Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd-graph.cgi?c=MOB24
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volkert Brendel, Stanford University,
http://grenlin1.zool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MXP5 and the 3' clone is MSD24.

COMMENT

Location/Qualifiers
1..79706
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOB24"
/clone_lib="Mitsui P1"
2398..3399
/note="gene_id:MOB24.1
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01995.1"
/db_xref="GI:9294038"
/translation="MGDSDEDTGTPKPKYKYSLNQRNHPMYSRIPKPHAYTNEDEDE
VEGEDPPGGYIRGNRFKQKPKYVSFEFAGSPDAKVADWRQEAFLVLEW
GDFPLQGRSLRNEDNWEVAEKSEELRMEKSFQCRRMIDDLKRYKPKIKYKES
GIGSKSFFKNLMLCVSPKSDGLAGGVDSGFVMTKVKYLDKNGFDMWDS
GDSDEDEDEVEYERKYNDAASYKMLADSVRFKGYKMEKSKREKMELEKMR
ADFQRLQKQIVDRAQSEIAREEENHHGGDDDEDEDEDEDEDEDEDEDE
complement(join(3794..4276,4363..4490,4577..4673,
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
8162..8217,8290..8385,8470..8998,9122..9462))
/note="gene_id:MOB24.2"
/codon_start=1
/evidence=not_experimental
/product="DNA repair protein MutS"

FEATURES
source

1..79706

CDS

1..79706
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOB24"
/clone_lib="Mitsui P1"
2398..3399
/note="gene_id:MOB24.1
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01995.1"
/db_xref="GI:9294038"
/translation="MGDSDEDTGTPKPKYKYSLNQRNHPMYSRIPKPHAYTNEDEDE
VEGEDPPGGYIRGNRFKQKPKYVSFEFAGSPDAKVADWRQEAFLVLEW
GDFPLQGRSLRNEDNWEVAEKSEELRMEKSFQCRRMIDDLKRYKPKIKYKES
GIGSKSFFKNLMLCVSPKSDGLAGGVDSGFVMTKVKYLDKNGFDMWDS
GDSDEDEDEVEYERKYNDAASYKMLADSVRFKGYKMEKSKREKMELEKMR
ADFQRLQKQIVDRAQSEIAREEENHHGGDDDEDEDEDEDEDEDEDEDE
complement(join(3794..4276,4363..4490,4577..4673,
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
8162..8217,8290..8385,8470..8998,9122..9462))
/note="gene_id:MOB24.2"
/codon_start=1
/evidence=not_experimental
/product="DNA repair protein MutS"

```

/protein_id="BAB01996.1"
/db_xref="GI:9294039"
/translation="MQQRSLSPFFQKPTAATKGLVSGDAASGGGGGGRFNVKKG
DAKDAVRAVSDEVSQITDTPPEKPRVPLPSGFKPAESAGDASSLESNTMHKF
KYVDRDCSERSREDVPLNDSLCMAKRVLEDMTPKEDKVPVLDNKRKLMDQPV
RSVEDIGVGGVPETPGMPKPRASRLKRVLEDMTPKEDKVPVLDNKRKLMDQPV
GGEKVEKNTKFKSLPSSRLDANRRPPDLYDKTLHPDPPVFKMSAQQKWS
VKSEYDIVLFFKVKYEIYELDAELGHEKLDNKMKGWGKRGVIGSGIDEAV
QKLARGTKVRIEQLSTDAKARGANTIIIPRLVQVLTPTSTSEGNIGDPAVHLLA
IKELKQKQSTVYGFADVCAALRFVWGSISDDASCAALGALIMVSPKEVLYDSK
GLSEAKLRKLYTIGSTAVQAPVQVPMGDTDAAGVNTIESNGYFKGSESWNCA
WLNCEGCVALSALGELINHLRLKLEDLVKGHLEFPYOVYRGCLRIDGQWVNLFI
NNSDCGPGSGTLYXVLCNVSPTGKRLILRNWICHLPLDVESINRKLVDVEFTANSE
MQITGQVLEKLPOLERLLGRKISSVRSASVLPALLKRVKQAKGQIVKGFPSG
IDLALQKQSNMSTLYKCKLPILVKGSGLELFLSOFEAADISDFPNYQNVQDTE
NAETLILIELFMTASRTQSEVHTIISCLWIRSFATAASISAGSMARVPYSEAT
DQNKTRGPIKIQGLMHPFAVADGQLPVPNDILLGEARSSGSIHPRSLLITGPNM
GKSTILRAICLAVIFAQIGCVPCSECEISLVITIFRLGASDRIMTGSTFIVECT
KFASVLONATQDSIVILDELGRGTSYTDGVAIAYSVFHLVEKVQCMLEATHVPLT
KEFASHPRVTSKHWACAFKRSYQPGCCDDLYFLRLTEGACPESGLQVALMAGI
PNQVETASGAQAKMSIGENFASSELRSSEFSLHEDWLKSLVGISRVAHNPAGE
DYDTFLCLMHEIKKSCVPR"
10131..10577
/note="gene_id:MOB24.3"
/codon_start=1
/evidence=not_experimental
/product="ethylene-responsive transcriptional
coactivator-like protein"
/protein_id="BAB01997.1"
/db_xref="GI:9294040"
/translation="MPSRYPGAVTDQWPEPVLRKSKOKSDLRQPKAVNAALRGVAV
QTVKFKASNNKCKSTAVPNTIKLEETEPRAMDYKAEVRLMIQAKLEKMSQ
ADLAKQINERQVQVEYENKAVPNQAVLAKERVGLVKGKIGK"
complement(join(11108..11361,11467..11528,11699..11836,
11934..12071,12161..12334,12421..12510,12718..13104,
13623..13765,14306..14405))
/note="gene_id:MOB24.4"
/codon_start=1
/evidence=not_experimental
/product="aldehyde dehydrogenase"
/protein_id="BAB01998.1"
/db_xref="GI:9294041"
/translation="MENGKNCATYVLPKPKTKLFINGOFIDAASQKTFETIDPRN
GEVTAIAGDKEDVDLAVNARAFBQWPRMTGPERAKLINKAEADLLEENLELA
KLDAVDGKLPQLGKYADIPATAGHPRYNAGAANKIGETIKMFRQSLFGTYLKEPIG
VYGNLFWNPSIFAFKAPMAAGCTMVVPAEQTSALFAYHLSKEAGIDEGVL
NIVFGSTAGAAASHMDVXVSTGTDVGRKIMQAAAASNLKYSLELKGKSPLL
IFNDADKADKALLGCFYMKGEICVASSRVFVQEGYIDVVKLEKAKDWTVGDP
FDPARQGPQDKRQFKEILSYIEHKNEGATLLTGGAIGDKGYFIQPTIFADVTE
MKIYDIFGVPVMSLMKFTVEEIKCANNTKYLGAAGILSQDIDLINTVRSIKAGI
IWNVYFGLDCPYGCKKSGNCRSGMDALDNLQTKSVYVPLHNSFW"
complement(join(19742..19933,20400..20657))
/note="gb|AAD25142.1"
/gene_id:MOB24.5
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01999.1"
/db_xref="GI:9294042"
/translation="MTPLSTASPPWPLLSLPSLSSSLNRAAPSFSPSRINHGPF
SCSPIAVQFALSRTKRLQVAMAPPEEKLIRNPLDFFIEWRPPGRPPDIFPK
FSPMKSLPPPPDYPDPEEEEEKKEEEDDDPEKEDQPDNRS"
join(23402..23459,23819..24021)
/note="gene_id:MOB24.6"
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02000.1"
/db_xref="GI:9294043"
/translation="MKNASLKPLLIIFLIVTSNLGAERKLTGVADVIVGEAASPO
YGIIDENDLNKSPCKRDIQTCRRGGFCNLILQKECI"
complement(join(25945..26640,26714..26824,26914..27012,
27119..27224,27291..27360,27436..27583))

```

CDS

CDS

CDS

CDS

CDS

```

/note="contains similarity to ubiquitin conjugating
enzyme"
/gene_id:MOB24.7"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02001.1"
/db_xref="GI:9294044"
/translation="MAQARLSRMQKELKLLSDPPHSGASPHLSAANGSGDFTF
STIDAIEGPEDEVYAGFINIKIQTIPERYFPPTVSPATPIYHNINSGRICLDI
LNLPKGAWQPSLINTSTVLSMRLISEPDDGLMCEVRSYKYNQFQKAREMT
EKYAVKADQSLTIQKNEGKSGESGNVGLKLTVPSSLSIAHTVKTREKAGDEQ
EDGKRAKAVQFGEGNSFENDGILKSRKLSLAPLSQSKKDLQCEBELIKVCSAC
KENKPNLNGKLSLGLQPCNDTLASPRISAASNNRSLKLSRLSLGELNEVSK
PEVLAQTMKLEMNQNDARSERGFENSVLSETSMAESIVLVDSDSDEEERVS
SRSLSLAKRVLKCRP"
28636..30168
/note="emb|CAB79678.1"
similar to unknown protein"
/gene_id:MOB24.8
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02002.1"
/db_xref="GI:9294045"
/translation="MSLIAYLSLFSNFSPLKLSRLSLNLLVLTSTNKKRWCL
ALPIYHLHATPLKSSITMLRFSQVPEPTPYREISLSRAFILTOAMSHGETLLWD
TISRLGLTRILMFSSASPHGQKQSTPTSPMDASLSSESPNRRHDLHLSEGL
GLQVLGVLFGLPYRMPISKDDPLKLPKPSDLNLSISFSQSLYELLASDRGH
TNNELVLITLWLNKSLRITLPLVSSRTFTVFGICLKNNGIMFTPTSRWDLGHSATI
CISPTHLGISFKQSCISSRKABIRISKVENPLFEIITEPSHALVAGLAKIGIMI
TSRSGDYORFNLSHSPLETFEKKRNIKFSTHDIQSLKNGIMISSPWSDDYR
SPFNLLDITPIVHLLGDEQLFLLVSARTNLPEYSSRLTAVNLLSDSFVED
PSTRDLTCAIMPYSLLKALMNFSEFIYHLLVGNASPNVFLNFIYFESLLFLY
LI"
Join(32831..33058,33143..33907)
/note="contains similarity to heat shock transcription
factor"
/gene_id:MOB24.9"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02003.1"
/db_xref="GI:9294046"
/translation="MEDDNNNNNNNVITVKTIONVNDPSTDLTLTWGPAHNSP
IIVVDLPDSORTLPAYFKHNDSVFRVOLTGFYSKVDPRDFRANEHLRQKHLIN
NIARSKHARGMQLDEGEIVREIRKEBORELPAEIQMRNRREIEKEKPEQMA
FLYAVVEDPDLPLRMLEKERKQQQVSDKKRRVTNSTVKSSEEEVEDEGVRVRY
MSSTSFSSSTENLYNHPDGWIVPMTOGQFSSTETGLVAKSMLSNSSSSSILTS
TFSLPEYVNGGGGCGSTQGERRYKETATFGVWESNPPTPPYFSLFRGCF"
complement(join(37054..37590,37670..37809,37894..38197,
38288..38398,38486..38584,38662..38772,39019..39066,
39311..39367))
/note="gene_id:MOB24.10"
unknown protein"
/codon_start=1

```

Query Match	Best Local Similarity	Score	446.6	DB 8	Length	79706			
Matches	853	Conservative	0	Mismatches	224	Indels	212	Gaps	6
QY	1	ATGTCTCTCGGCGCGTCTCCGGGACTGGTTCCTCCATCTCCACCAT	-----CA 51						
DB	51787	ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCAACCATCTCCACCGTCTCTCCCA	51846						
QY	52	AACTGCACACACACACTCTCTCCAGCTTCGCTCCCTCCCTCCACCA	-----CA 102						
DB	51847	AACTGCACACACACACTCTCTCCAGCTTCGCTTCCTCTCCCTCCACCACTCT	51906						
QY	103	CCTCTCTCTCTCCCGCGCATCCCTATCTCCGAGATCTCCCTCTCTCTCTCTCTCT	162						
DB	51907	TCCTCTCTCTCTCCCGCGCATCCCTATCTCCGAGATCTCCCTCTCTCTCTCTCT	51966						
QY	163	ACACCTTCT	222						
DB	51967	CCT	52014						

QY 1101 AATGGAAATGGACACAGATTCAGAAATGCTCTTGGATGCTGCTTAAGGACCTTCTTATCT 1160
DB 1158 TTGATGTTGGCTTACAAAGAGTGAAGATTCATGCTAGGATGAGGCTTGCATATTT 1217
QY 1161 TCATGAAGATTGCA-----ATCCTAAATCAT 1187
DB 1218 GCATGAAGACTGAAGAAATCTTTACTCACAATTTGCAATCAGTCACCTCGCATAT 1277
QY 1188 TCACCGTGATATCAAGGCTTCAACATATGATAGATTTCAAGTTTGAAGCTTAAGGTGTC 1247
DB 1278 CCACAGATATCAAGAGCTGCAACATCTTCTTGATTCAGTTTGGACCAAGGTGTC 1337
QY 1248 TGATTTTGGCTTGTGAAGTTCGCTGATACAAACAGCGCATGATACACAGCTGTGAT 1307
DB 1338 AGATTTGGATGGCTAAGCTATCTCAAGACAATATATCATGTCTCCACTCGGTCTAT 1397
QY 1308 GGAACCTTTGGTACTTGGCTCCGGAATACGCTCAAGCGGAAGCTCACGGAGAATC 1367
DB 1398 GGAACCTTTGGATCTAGCTCCAGAGTATGCTCAAGCGGAAGTATTCGACAAATC 1457
QY 1368 TGACGTTTCTCATTTGGCTTGTGCTTTGGAGCTCAATTAAGTCAAGCTCGACCGCTTGA 1427
DB 1458 TGATGTTTCTCATTTGGATTAATGCTTCTGAGCTCAATACCGGAAGACCTCTCTGGA 1517
QY 1428 TGCCAAACATGCTATGATGACAGCTTACTGCTGGGACAGCAATGCTTAAACCG 1487
DB 1518 T---CTAACTGAGAAATGGAAGATGCTTGGTAGTGGCAAGGCTTTGTGTTGAA 1574
QY 1488 AGCATCTGACGAAGAGACTTTGAGGCTTACTGATGCAAGATGAATAAGGGTATCA 1547
DB 1575 AGCAGCTCAAGATGAGATTAACAACCAATGCTGATCCAGCTAGAGCTAAACTACAG 1634
QY 1548 CAGAGAGAGATGGCTCCAGGTTGCTTGCTGCTCGGCTTGTGCTGCTTCAAGCTAGCTCG 1607
DB 1635 TCATCAAGAGATGGTTCAATGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694
QY 1608 CCGAGAGCTCGCATGACCCAG-----ATTGT 1634
DB 1695 AAGCGGCTTAAATGAGCCAGGTTCAAAAATCATACCACTTGTGCTTATTTATTT 1754
QY 1635 GCSTGCTTAGAAGAAATGATCATCTGCTGATCTTAACGAAGGATGAGACCAAGTCA 1694
DB 1755 ACGAGCTAGAGAGATGCTCATGATGATCTAGTGGAGAACAGACCAAGGACA 1814
QY 1695 AAGCAATGATACAGCTCATACGGAGAACACCGATTAAGTACTCGAGCCAGCAATGA 1754
DB 1815 AAGCAGCTACTTGAAGCCCGGAGCGTCAAGTATGACGAAGCTCGTACAGGSC 1874
QY 1755 AGCATGAAGAGTTTAGGAATGCGACTTGGACTCAAGAGTACAGCCACCGGTGA 1814
DB 1875 AGCATGAAGAGTTTAGGAATGCGCTTAGAGAAATGAAGAAATCAAGCAGTGA 1930
QY 1815 GTACAGTAATCCGACAGTACTGACTGTACCCGCTGCTGCTTCAAGCAGCGAGGCA 1874
DB 1931 --AATGTTGGAACAGTGAATGGCTTAAACCCCTCTGCTTCAAGTAGTGAAGAT 1988
QY 1875 AACCACACCGGAATGGA 1892
DB 1989 GAATAGAGGCTCAATGA 2006

RESULT 13
AX088884
LOCUS AX088884
DEFINITION Sequence 9 from Patent WO0114563.
ACCESSION AX088884
VERSION AX088884.1 GI:3397647
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2104)
Goring,D. and Silva,N.
Proline-rich extensin-like receptor kinases
Patent: WO 0114563-A 9 01-MAR-2001;
Goring, Daphne (CA); Silva, Nancy (CA)
FEATURES
Location/Qualifiers
1..2104
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 609 a 517 c 455 g 523 t
ORIGIN
Query Match 22.0%; Score 427.2; DB 6; Length 2104;
Best Local Similarity 61.8%; Pred. No. 9.1e-76;
Matches 790; Conservative 0; Mismatches 413; Indels 75; Gaps 4;
QY 681 TCACACACCGCTTCATGAGACGACGCGGCTCGGACTACTCGGACCGTCCAGTCT 740
DB 774 TCACACGACCACTGTCTGTCTCTCACAGTATGCTTCCAACTTAACCGTCCGAACTGC 833
QY 741 TCCTCCACCGCTTCCAGGGCTTGTGTAGGCTTCTCCAAAAGACACTTTCACATACGAGGA 800
DB 834 TATACCGTGCCTCAGCTGCAACTCTTGTCACACCAAGACACTTTCACATACGATGA 893
QY 801 GCTAGTAGAGCCCAATGGTTTCTCCGAGGGGAACTTGTAGGACAAAGCGGTTGCG 860
DB 894 ACTGTCCATTGCAACAGAGGTTTGGCTCAGTCAAAATTTGCTAGGACAAAGGATTTGG 953
QY 861 TTACGTGCAACAAGAGTGTGTGCTAGTGGGAAAGAGTTCTGTGTAACAGTGAAGT 920
DB 954 GTATGTTCTAAAGAGGTTTGGCTAGTGGCAAGAGTTGCTAGTGAAGAGTCTTAACT 1013
QY 921 TGGGAGTGTTCAGGAGAGAGGAGTTTTCAGGCGAGGTTGAGATCATCAGCAGGTTCA 980
DB 1014 TGGAAATGACAGGAGGAGCGGAGTTTCAAGCAGAGGTTGATATCATTAAGCCGTGCCA 1073
QY 981 CCACAGCATCTGTGCTCTTGTGTTTATTCATCGCGGTGCCAAAAGATTCGTTGT 1040
DB 1074 TCATGTCATCTGTTTCTCTTGTGATTTTCATCTCTGTTGTTCAAGAGCTTTTGT 1133
QY 1041 CTATGAGTTTCTTCAACAACATCTCGAGCTTCCCTCCATCGGCGGAGGAGGCTTAC 1100
DB 1134 TTATGAGTTTATACCTTAACAACACTCTTGAATTTCTATTCATGGAAGGCTCGTCGGT 1193
QY 1101 AATGGAATGACCAACAGATTTGAAGATTTGCTCTGATCAAAACAGCATGTATCAACAGTGTGAT 1160
DB 1194 TTGTGATTGGCTTCAAGAGTGAAGATTCGATTGGGATCAGTAGAGGCTTGCATATT 1253
QY 1161 TCATGAAGATTGCA-----ATCCTAAATCAT 1187
DB 1254 GCATGAAGACTGTAAGAAATCTTTATCTACATATTTGCTCAGTCACCGCTCGCATAT 1313
QY 1188 TCACCGTGATATCAAGGCTTCAACATATTTAGATTTCAAGTTTGAAGCTTAAGGTGTC 1247
DB 1314 CCACAGAGATTCAAAGCTGCAACCAATCTTCTGATTTCTGTTTGAAGCAAGGTGTC 1373
QY 1248 TGATTTTGGCTTGTCTAGATTTCTCTGTATACAAACAGCATGTATCAACAGCTGTGAT 1307
DB 1374 AGATTTGATTGGCTTAAGCTATCTCAAGACACTATACATCTCTCCACTCGGCTCAT 1433
QY 1308 GGAACCTTTGGTACTTGGCTCCGGAATACGCTGCAAGCGGAAGACTCACGGAGAGTC 1367
DB 1434 GGAACCTTTGGATACTTAGCTCCAGAGTATGATCAAGCGGAAGAGTTATCCACAATC 1493
QY 1368 TGACGTTTCTCATTTGGGCTTGTGCTTTTGGAGCTCATTAAGTGGAGCTGACCCGTTGA 1427
DB 1494 TGATGTTTCTCATTTGGAGTAAATGCTTCTGAGCTCATTAACCGGAAGACCTCTCTGGA 1553
QY 1428 TGCCAAACATGCTATGATGACAGCTTACTGTTGCTGGCAGCAGCAATTCGTTAAACG 1487
DB 1554 T---CTAACTGGAAGAAATGGAAGATGCTTGGTAGATTGGCAAGGCTTTGTGTTGGA 1610

LOCUS AX088887 2261 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 12 from Patent WO0114563.
ACCESSION AX088887
VERSION AX088887.1 GI:13397650
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2261)
Goring, D. and Silva, N.
TITLE Proline-rich extensin-like receptor kinases
JOURNAL Patent: WO 0114563-A 12 01-MAR-2001;
Goring, Daphne (CA); Silva, Nancy (CA)
FEATURES
source
1..2261
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 567 a 652 c 456 g 586 t
ORIGIN

Query Match 21.3%; Score 413.6; DB 6; Length 2261;
Best Local Similarity 67.0%; Pred. No. 5e-73;
Matches 606; Conservative 0; Mismatches 289; Indels 9; Gaps 1;

QY 782 GCACCTTTTCACATACGAGGAGTAGCTAGAGCCACCACCAATGGTTTCTCCAGCGGAACCTTGT 841
DB 1151 GAATGTTCTTCCACGAGACACTTCAAGACCACTGGTGGATTTTCAGAGGAGAACCTTT 1210

QY 842 TAGGACAGCGGGTTCGGTTAGTGCACAAAGGTGTGTTGCCCTAGTGGGAAGAAGTTG 901
DB 1211 TGGGAGAGCGGTTTCGGATATGTTCAAGAGGAGTGTGAAAGACGGGACAGAAAGTTG 1270

QY 902 CTGTGAACAGTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTCAGGACAGAGGTTG 961
DB 1271 CGGTGAACAGCTGAAGATGGGAGCTATCAAGGGGAAGAGATTCGAAGCTGAGGTTG 1330

QY 962 AGATCATCAGCAGAGTACCAAGGATCTGCTGCTCTTGTGTTTATTTGATTCGCTCG 1021
DB 1331 ACACATCAGTGGGTTTCATCAATGACACCTGCTTCATGTTGTTTATTTGCTTAATG 1390

QY 1022 GTGCCAAAGATTTGTTGCTATGATTTCTTCTTACACAAATCTCGAGCTTCACCTCC 1081
DB 1391 GAGATAAAGACTCTTGGTTTACGAGTTTCTTCTTAAAGATACCTTGGAGTTCCACTTC 1450

QY 1082 ATGGCGAGGACGGCTACAAATGGAATGGAGCACCAGATGAAGATGCTCTGGATCTG 1141
DB 1451 ATGAGAACAGAGGAGGCGTGTGGATGGGAATGAGGCTCAGGATTCCTGTAGGAGAG 1510

QY 1142 CTAAGGACTTTCTTATCTCATGAGATTCGAATCCTAAATCAATTCACCGTGATATCA 1201
DB 1511 CAAAGGATTAGCTTATCTCATGAGATTCGAGTCCCACTCAATATCAACGCTGATATCA 1570

QY 1202 AGGCTTCAAACATATGATAGATTTCAAGTTTGAAGCTAGGTTGCTGATTTTGGTCTTG 1261
DB 1571 AAGCAGCTAATATCTTCTAGATTCACAAATTTGAGCAAAAGTCTCTGACTTTGGACTAG 1630

QY 1262 CTAGATGCTTCTGATACAAA -----CAGCATGTATCAACAGCTGTGATGGAA 1312
DB 1631 CCAAGTTTTTCACAGACCAAAATTCATTCATCATATATCTCTACTCGAGTGTAGGAA 1690

QY 1313 CCTTTGGTACTTGGCTCGGAATACGCTGCAAGGAAAGCTCACGGAAGTCTGACG 1372
DB 1691 CTTTCGGATATAGCTCCAGAAATACGGTCCAGTGGTAAAGTAACTGATTAATCAGATG 1750

QY 1373 TTTTCTCATTTGGGTTGTGTTTGGAGCTCATTTACTGGACGTCGACCCCTTGTATGCCA 1432
DB 1751 TATATTCCTTTGGGTCGTGTTCTTCTAGAACTCATCACTGGAGCTCCATCAATTTTCGCCA 1810

QY 1433 ACAATGCTATAGATCAGACGCTTAGTTGACTGGGACAGCAATGCTTAAACGAGCAT 1492
DB 1811 AAGATCTTCCACAACAGAGTTTAGTAGACTGGGAGGCCAATGCTTACGAAGCA 1870

Search completed: July 7, 2003, 09:38:57
Job time : 5136 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 06:45:11 : Search time 465 seconds
(without alignments)
94.4.866 Million cell updates/sec

Title: US-10-086-464-1

Perfect score: 1944

Sequence: 1 atgctctgcgcgcgtctcc.....atagtgacattctctttaa 1944

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1944	100.0	1944	22	Brassica napus PER
2	1341.2	69.0	2324	21	Arabidopsis thalia
3	215.8	11.1	1703	21	zea mays DNA fragm
4	204.6	10.5	1419	21	Arabidopsis thalia
5	199	10.2	1353	21	Arabidopsis thalia
6	199	10.2	1626	21	Arabidopsis thalia
7	197.2	10.1	1455	21	Arabidopsis thalia
8	197.2	10.1	1751	21	Arabidopsis thalia
9	197	10.1	1569	21	Arabidopsis thalia

10	194.8	10.0	1479	21	AAC43040	Arabidopsis thalia
11	194.6	10.0	1674	21	AAC39626	Arabidopsis thalia
12	192.4	9.9	3108	21	AAAT9734	Eucalyptus grandis
13	192.2	9.9	2820	22	AAAT7057	Arabidopsis gene #
14	192	9.9	2520	22	AAAT7096	Arabidopsis thalia
15	191.8	9.9	1848	21	AAC50527	A thaliana recepto
16	191.8	9.9	1910	24	AAAT45908	Arabidopsis thalia
17	190.6	9.8	1572	21	AAC42898	Arabidopsis thalia
18	190.6	9.8	1920	24	AAAT45905	A thaliana recepto
19	190.6	9.8	2012	21	AAC44847	Arabidopsis thalia
20	186.8	9.6	990	21	AAC41702	Arabidopsis thalia
21	186.2	9.6	1563	21	AAC43257	Arabidopsis thalia
22	186.2	9.6	2011	22	AAAT2557	Arabidopsis thalia
23	186	9.6	1887	21	AAC42821	P. patens protein
24	186	9.6	1949	24	AAAT45904	Arabidopsis thalia
25	184.2	9.5	1587	24	AAAT45901	A thaliana recepto
26	184.2	9.5	2383	21	AAC49600	Arabidopsis thalia
27	182.6	9.4	2384	21	AAC37339	Arabidopsis thalia
28	182.2	9.4	1329	21	AAC36224	Arabidopsis thalia
29	182.2	9.4	1651	21	AAC36587	Arabidopsis thalia
30	182	9.4	2868	21	AAAT9266	Arabidopsis thalia
31	179.4	9.2	1449	21	AAC43051	pinus radiata cell
32	179.4	9.2	1506	21	AAC49445	Arabidopsis thalia
33	179.4	9.2	1757	21	AAC41794	Arabidopsis thalia
34	178.6	9.2	1920	24	AAAT45903	Arabidopsis thalia
35	178.4	9.2	3239	21	AAAT9271	A thaliana recepto
36	177.8	9.1	1814	19	AAAT06571	Eucalyptus grandis
37	177.4	9.1	3128	21	AAAT46908	Daucus carota SERK
38	176.8	9.1	2217	21	AAC40913	Arabidopsis thalia
39	176.4	9.1	1740	24	AAAT45899	Arabidopsis thalia
40	175	9.0	1575	21	AAC43256	A thaliana recepto
41	175	9.0	1839	21	AAC50528	Arabidopsis thalia
42	175	9.0	1925	24	AAAT45907	Arabidopsis thalia
43	174.6	9.0	2432	21	AAAT9263	A thaliana recepto
44	174.2	9.0	1555	21	AAC37541	Pinus radiata cell
45	173.2	8.9	2087	24	AAAT45897	Arabidopsis thalia
						A thaliana recepto

ALIGNMENTS

RESULT 1

AAAT77094

ID AAAT77094 standard; DNA; 1944 BP.

XX AAAT77094;

XX AAAT77094;

XX 17-MAR-2001 (first entry)

XX Brassica napus PERK1 DNA.

XX Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.

XX Brassica napus.

XX WO200114563-Al.

XX 01-MAR-2001.

XX 18-AUG-2000; 2000WO-CA00966.

XX 19-AUG-1999; 99US-0149466.

XX 13-OCT-1999; 99US-0159122.

XX (GORI/) GORING D.

XX (SILV/) SILVA N.

XX Goring D, Silva N;

XX WPI; 2001-244305/25.

XX New proline-rich, extensin-like receptor kinase nucleic acids and

PT polypeptides useful for increasing plant wounding or pathogen
 PT resistance, or for producing transgenic plants with increased wounding
 PT or pathogen resistance -

XX Claim 6; Fig 1; 91pp; English.

XX The present invention relates to proline-rich extensin-like
 CC receptor kinase (PERK). The PERK nucleic acids and polypeptides
 CC are useful for increasing the resistance of plants to wounding
 CC and pathogens. These are also useful for producing transgenic
 CC plants with increased wounding and pathogen resistance compared
 CC with a wild type plant, as well as in assays for identifying
 CC and developing compounds to inhibit and/or enhance polypeptide
 CC function directly.

XX Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 other;

Query Match 100.04; Score 1944; DB 22; Length 1944;
 Best Local Similarity 100.04; Pred. No. 0;
 Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTCCTCGGCGCCGCTCCGGGAGCTGGTTCGGCCCTCATCTCCACCATCAACATCCACA	50
Db	1	ATGTCCTCGGCGCCGCTCCGGGAGCTGGTTCGGCCCTCATCTCCACCATCAACATCCACA	50
QY	61	ACCACCATCTCTCCAGCTTCGCGCTCCCTCCACACACACTTCTCTCTCCGCGG	120
Db	61	ACCACCATCTCTCCAGCTTCGCGCTCCCTCCACACACACTTCTCTCTCCGCGG	120
QY	121	CCATCCCATATCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180
Db	121	CCATCCCATATCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180
QY	181	CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCT	240
Db	181	CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCT	240
QY	241	CCCGCTCCCACTAGCGCGGATCTCCACCGGACCTGTTACTCTCTCTCTCTCTCTCTCT	300
Db	241	CCCGCTCCCACTAGCGCGGATCTCCACCGGACCTGTTACTCTCTCTCTCTCTCTCTCT	300
QY	301	CCACTCTCAGTCCGAGGACCACTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
Db	301	CCACTCTCAGTCCGAGGACCACTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
QY	361	CCATCT	420
Db	361	CCATCT	420
QY	421	GGATCT	480
Db	421	GGATCT	480
QY	481	AAGAAGAAACGAGGAG	540
Db	481	AAGAAGAAACGAGGAG	540
QY	541	GGTCCCAAGCCGAGGAG	600
Db	541	GGTCCCAAGCCGAGGAG	600
QY	601	ACACCAAGCTCAGATCATCT	660
Db	601	ACACCAAGCTCAGATCATCT	660
QY	661	CCAGGCAAGCT	720
Db	661	CCAGGCAAGCT	720
QY	721	TACTCGGACGCT	780
Db	721	TACTCGGACGCT	780

QY	781	AGCATTTCACATACGAGGAGCTAGCTAGAGCCACCACCAATGGTTTCTCCGAGCGCACTTG	840
Db	781	AGCATTTCACATACGAGGAGCTAGCTAGAGCCACCACCAATGGTTTCTCCGAGCGCACTTG	840
QY	841	TTAGGACAAAGCGGGTTCGGTTCAGTGCACAAAGGCTGTGTCTAGTGGGAAAGAGTT	900
Db	841	TTAGGACAAAGCGGGTTCGGTTCAGTGCACAAAGGCTGTGTCTAGTGGGAAAGAGTT	900
QY	901	GCTGTGAAGCAGTTGAAGTTGGAGTTGGTGTTCAGGAGAGAGAGAGTTTCAGGAGAGGTT	960
Db	901	GCTGTGAAGCAGTTGAAGTTGGAGTTGGTGTTCAGGAGAGAGAGAGTTTCAGGAGAGGTT	960
QY	961	GAGATCATCAGCAGAGTTCCACACAGGCTATCTGGTGTCTCTCTCTCTCTCTCTCTCTCT	1020
Db	961	GAGATCATCAGCAGAGTTCCACACAGGCTATCTGGTGTCTCTCTCTCTCTCTCTCTCTCT	1020
QY	1021	GGTCCCAAAAGATTGCTGTCTATGAGTTTGTCTTAACAAACATCTCCAGGCTTCACCTC	1080
Db	1021	GGTCCCAAAAGATTGCTGTCTATGAGTTTGTCTTAACAAACATCTCCAGGCTTCACCTC	1080
QY	1081	CATGGCGAGGAGCGGCTACAAATGGAATGGAGATTCGAGCAGAGATTCGCTCTTGGATCT	1140
Db	1081	CATGGCGAGGAGCGGCTACAAATGGAATGGAGATTCGAGCAGAGATTCGCTCTTGGATCT	1140
QY	1141	GCTAAGGAGCTTCTTATCTTCATGAGATTGCAATCTTAAATCATTCACCGTGATC	1200
Db	1141	GCTAAGGAGCTTCTTATCTTCATGAGATTGCAATCTTAAATCATTCACCGTGATC	1200
QY	1201	AAGGCTTCAACATATTTAGATTTCAAGTTTGAAGCTTGAAGCTTGAAGTTTGTCTT	1260
Db	1201	AAGGCTTCAACATATTTAGATTTCAAGTTTGAAGCTTGAAGCTTGAAGTTTGTCTT	1260
QY	1261	GCTAAGATTGCTTCTGATCAACAGCGCATGATCAACAGCGTGTGATGGAGACCTTTGGG	1320
Db	1261	GCTAAGATTGCTTCTGATCAACAGCGCATGATCAACAGCGTGTGATGGAGACCTTTGGG	1320
QY	1321	TACTTGGCTCCGGAATACGCTCGAAGGAGCTACCGGAGAGTCTGAGCTTCTCTCA	1380
Db	1321	TACTTGGCTCCGGAATACGCTCGAAGGAGCTACCGGAGAGTCTGAGCTTCTCTCA	1380
QY	1381	TTTGGCTTGGCTTTTGGAGCTCATTTACTGACGCTGACCGCTTGTATGCCAACATGTC	1440
Db	1381	TTTGGCTTGGCTTTTGGAGCTCATTTACTGACGCTGACCGCTTGTATGCCAACATGTC	1440
QY	1441	TATCTAGATGACAGCTTATGTTGAGCTCATTTACTGACGCTGACCGCTTGTATGCCAACATGTC	1500
Db	1441	TATCTAGATGACAGCTTATGTTGAGCTCATTTACTGACGCTGACCGCTTGTATGCCAACATGTC	1500
QY	1501	GGAGACTTTGAGGGTTTACCTGATGCAAGATGAAATATGGTATGACAGAGAGATG	1560
Db	1501	GGAGACTTTGAGGGTTTACCTGATGCAAGATGAAATATGGTATGACAGAGAGATG	1560
QY	1561	GCTCGATGTTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1620
Db	1561	GCTCGATGTTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1620
QY	1621	ATGAGCCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
Db	1621	ATGAGCCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
QY	1681	ATGAGCCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1740
Db	1681	ATGAGCCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1740
QY	1741	AGCCAGTACATCAAGACATGAAGAGTTTGAAGAAATGGCACTTGAACCTCAAGAGTAC	1800
Db	1741	AGCCAGTACATCAAGACATGAAGAGTTTGAAGAAATGGCACTTGAACCTCAAGAGTAC	1800
QY	1801	AAGCCAGGCTGAGTACATCAATCCGACAGTATGATGATGATGATGATGATGATGATGATG	1860
Db	1801	AAGCCAGGCTGAGTACATCAATCCGACAGTATGATGATGATGATGATGATGATGATGATG	1860
QY	1861	AGCAGCGAGGGCCAAACACACCGGAATGGAGATGGGAGATTAGAGAAACCGCTCAG	1920

```
Db      1861 1861 AGCAGGAGGCGCAACACACGCGAAATGGAGATGGGAGATTAAGAGAACCGGTACG 1920
QY      1921 1921 GGTATAGTGGACTTCTCTTTAA 1944
Db      1921 1921 GGTATAGTGGACTTCTCTTTAA 1944

RESULT 2
AAC36968
ID      AAC36968 standard; DNA; 2324 BP.
XX
AC      AAC36968;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 15710.
XX
KW      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121325.
PR      05-MAR-1999; 99US-0123180.
PR      03-MAR-1999; 99US-0123348.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131445.
PR      30-APR-1999; 99US-0132048.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132484.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134221.
PR      16-MAY-1999; 99US-0134370.
PR      19-MAY-1999; 99US-0134768.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      28-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142154.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145919.
PR      02-AUG-1999; 99US-0145921.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
```



```
Db 1272 TATCTTCATGAGATGCAACCGGAAATCATTCACCGTGATTAAGGCGTCAACATA 1331
      |||
QY 1216 TTGATAGATTCAAGTTTGAAGCTTGAAGTTGCTGATTTGGTCTTGTAAAGATTGCTTCT 1275
      |||
Db 1332 TTGATTGATTTCAAAATTTGAAGCTTGAAGTTGCTGACTTTGGTCTTCCAGGATGCTTCT 1391
      |||
QY 1276 GATACAAACACCATGTTATCACACAGTGTGATGGGAACCTTTGGGTACTTGGTCCGGAA 1335
      |||
Db 1392 GATACAAACACCATGTTATCACACAGTGTGATGGGAACCTTTGGGTACTTGGTCCGGAA 1451
      |||
QY 1336 TACGCTGCAAGCGGAAAGCTCACGGAGAGTGTGAGCTTTTCTCANTTGGCTTGGCTT 1395
      |||
Db 1452 TATGCTGCAAGCGGAAAGCTCACGGAGAGTGTGAGCTTTTCTCANTTGGCTTGGCTT 1511
      |||
QY 1396 TTGGAGCTCATTTACCTGGAGCTGCAAGCCCTTGTATGCCAACAATGTCTATGATGACAGC 1455
      |||
Db 1512 TTGGAGCTCATTTACCTGGAGCTGCAAGCCCTTGTATGCCAACAATGTCTATGATGACAGC 1571
      |||
QY 1456 TTAGTTGACTGGCAGCAGCATTGCTTACCGAGCACTGAGCAAGGAGACTTTGAGGCT 1515
      |||
Db 1572 TTAGTTGACTGGCAGCAGCATTGCTTACCGAGCACTGAGCAAGGAGACTTTGAGGCT 1631
      |||
QY 1516 TTAGCTGATGCAAGAGTGAATTAATGGTATGACAGAGAGTGGCTCGCATGCTTGGCT 1575
      |||
Db 1632 TTGGCTGATCAAGATGCTTAATGATAT-AGAGAGAGGATGGCTCGCATGCTTGGCT 1690
      |||
QY 1576 TGTGCTGGCGTTGTTGCGCATTCAGCTCGCCGAGACCTCGCAAGAGCCAGATTGTG 1635
      |||
Db 1691 TGGCTGCAAGCTTGTGTTGCGCATTCAGCTCGCCGAGACCTCGCAAGAGCCAGATTGTG 1750
      |||
QY 1636 CGTGGGTTAGAGGAATGTTATCACTGTGATGATGATGATGATGATGATGATGATGATGAT 1695
      |||
Db 1751 CGGGCTTAGAGGAATGTTATCACTGTGATGATGATGATGATGATGATGATGATGATGAT 1810
      |||
QY 1696 AGCAATGATATCAGCTATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1755
      |||
Db 1811 AGCAACGTATACAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1870
      |||
QY 1756 GACATGAGAGAGTTAGAGAAATGCACTTGGAACTCAAGAGTACAAAGCCAGGCTGAG 1815
      |||
Db 1871 GACATGAGAGAGTTAGAGAAATGCACTTGGAACTCAAGAGTACAAAGCCAGGCTGAG 1930
      |||
QY 1816 TACAGTATCCAGCAGTACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1875
      |||
Db 1931 TACAGTATCCAGCAGTACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1990
      |||
QY 1876 ACCACAGCGAAATGGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1935
      |||
Db 1991 GCCACAGGAAATGGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2050
      |||
QY 1936 TCTCTTTAA 1944
      |||
Db 2051 TCTCTTTAA 2059
      |||

RESULT 3
AAC46629
ID AAC46629 standard; DNA; 1703 BP.
XX
AC AAC46629;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 50839.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; cont; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
```

```
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151436.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.1%; Score 215.8; DB 21; Length 1703;
Best Local Similarity 59.0%; Pred. No. 1.6e-41;
Matches 430; Conservative 0; Mismatches 287; Indels 12; Gaps 3;

QY 769 GGCCTTCCTCAAAAGACATTCACATACGAGGAGGTAGCTAGAGCCACCAATGGTTCTTCC 828
DB 209 GCCATCTCAGCCCAACATTCACCTCCGTGACCTGTGTCGCCACCAAGGACCTCAGG 268
QY 829 GAGCGCAACTTTTAGGACAAGCGGGTTCGGTTAGTGTCACAAAGGTG---TGTGCGCT 885
DB 269 CCCGAGTCTTCTCGGAGAGGAGGTTTGGACGTGTTTACAGGGCGCGCTTGAGAGC 328
QY 886 AGTGGGAAGAGTTGCTGTGAGCAGTTGAAGTTGGAGTGGAGTGGTTCAGGAGAGGAG 945
DB 329 ACAGCCAGGTGTGCTTATAAGCAGCTTAACAGGATGGCTTCAAGGAACAGAGAA 368
QY 946 TTTTCAGGAGGTTGAGATCATCAGCAGATTCACCAAGCATCTTGTGTCTTGT 1005
DB 389 TTTCTAGTAGAGTTCTCATGCTCAGCTTACTACATCATCAAACTGCTTATTTGAT 448
QY 1006 GGTATTCATGCGCGGTGCCAAAGATTTGCTGTATGATGTTTGTCTTCAACAAAT 1065
DB 449 GGTATTGCTGTATGGAGACCCAGAGACTTCTTGTATTGATATATGATGCTTCTGATCA 508
QY 1066 CTCGAGCTTCACCTCCATGGCGAGGACGGCTA-----CAATGAATGAGCACCA 1119
DB 509 CTTGAAGATCATTTTGCACGATCTACTCTTGAAGAGGCGCTTGCATGGAAACACAGG 568
QY 1120 TTGAAGATTGCTTCTTGGATCTCTTAAGGACATTTCTTATCTTCATGAAGATTGCAATCT 1179
DB 569 ATGAAGATTGAGCAGGTGCTGCCAAGAGCTGGAGTACCTTCATGACAAACCTATCA 528
QY 1180 AAATCATTCACCGTGATATCAAGGCTTCAACATATTGATAGATTCAAGTTTGAAGCT 1239
```

```
Db      629  CCAGTTATTATAGGGAATTCAGTCATCGAACAATCTGTGGTGAAGATTCACCCG 688
Qy      1240 AAGTTCTGATTTGGTCTTGAAGATTCCTCTGTATACAAACA---CGCATGTCNCA 1296
Db      589  AAGCTGCTGATTTGGACTGCTTAAGTTGGGTCAGTTGGCGAATAATCATCTGCTCA 748
Qy      1297  ACAGCTGTATGGGAACCTTTGGGTACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTC 1356
Db      749  ACAGCTGTATGGGTACATATGGTTATTGTGCACCAAGATATGCTATGACAGGCGAGCTG 808
Qy      1357  ACGGAGAGCTGACGCTTCTCATTTGGCGCTTGGCTTTTGGAGCTCATTTACTGGAGCT 1416
Db      809  ACAGTGAAGTCTGACGCTGATAGTTTGGGCTTCTGTCTGTAGAGTTGATTACTGGCCG 868
Qy      1417  CGACCCGCTTGTATGCAACAATGCTCATGTAGATCACAGCTTAGTTGATGGGCGACGCA 1476
Db      869  AGGCTATTGACACACACACACACACATGGAGAACAAATCTGTCTATGGGCGAGCTCT 928
Qy      1477  TTGCTTAAC 1485
Db      929  CTG-TTAAC 937

RESULT 4
AAC43029
ID      AAC43029 standard; DNA; 1419 BP.
AC      AAC43029;
XX
XX      17-OCT-2000 (first entry)
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 37767.
XX
XX      Hybridisation assay; genetic mapping; gene expression control;
XX      protein identification; signal transduction pathway;
XX      metabolic pathway; promoter; termination sequence; ss.
XX
XX      Arabidopsis thaliana.
XX      EPI033405-A2.
XX
XX      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-0301439.
XX
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
XX      09-MAR-1999; 99US-0123548.
XX      23-MAR-1999; 99US-0125788.
XX      25-MAR-1999; 99US-0126264.
XX      29-MAR-1999; 99US-0126785.
XX      01-APR-1999; 99US-0127462.
XX      06-APR-1999; 99US-0128234.
XX      08-APR-1999; 99US-0128714.
XX      16-APR-1999; 99US-0129845.
XX      19-APR-1999; 99US-0130077.
XX      21-APR-1999; 99US-0130449.
XX      23-APR-1999; 99US-0130510.
XX      28-APR-1999; 99US-0130891.
XX      30-APR-1999; 99US-0131449.
XX      30-APR-1999; 99US-0132048.
XX      30-APR-1999; 99US-0132407.
XX      04-MAY-1999; 99US-0132484.
XX      05-MAY-1999; 99US-0132485.
XX      06-MAY-1999; 99US-0132486.
XX      06-MAY-1999; 99US-0132487.
XX      07-MAY-1999; 99US-0132863.
XX      11-MAY-1999; 99US-0134256.
XX      14-MAY-1999; 99US-0134218.
XX      14-MAY-1999; 99US-0134219.
XX      14-MAY-1999; 99US-0134221.
XX      14-MAY-1999; 99US-0134370.
XX      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      25-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
```


KW Metabolic pathway; proutter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EF1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-C301439.
PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138054.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 15-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150565.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

Query Match	10.2%	Score 199	DB 21	Length 1353
Best Local Similarity	56.9%	Pred. No. 1.5e-37		
Matches 408	Conservative	0	Mismatches 300	Indels 9
Gaps	2			
PR 10-SEP-1999	990S-0153070.			
PR 13-SEP-1999	990S-0153758.			
PR 15-SEP-1999	990S-0154018.			
PR 16-SEP-1999	990S-0154039.			
PR 20-SEP-1999	990S-0154779.			
PR 22-SEP-1999	990S-0155139.			
PR 23-SEP-1999	990S-0153486.			
PR 24-SEP-1999	990S-0153659.			
PR 28-SEP-1999	990S-0156458.			
PR 29-SEP-1999	990S-0156596.			
PR 03-OCT-1999	990S-0157117.			
PR 05-OCT-1999	990S-0157753.			
PR 08-OCT-1999	990S-0157865.			
PR 07-OCT-1999	990S-0158029.			
PR 08-OCT-1999	990S-0158232.			
PR 12-OCT-1999	990S-0158369.			
PR 13-OCT-1999	990S-0159293.			
PR 13-OCT-1999	990S-0159294.			
PR 13-OCT-1999	990S-0159295.			
PR 14-OCT-1999	990S-0159329.			
PR 14-OCT-1999	990S-0159330.			
PR 14-OCT-1999	990S-0159331.			
PR 14-OCT-1999	990S-0159637.			
PR 18-OCT-1999	990S-0159838.			
PR 21-OCT-1999	990S-0159584.			
PR 21-OCT-1999	990S-0160741.			
PR 21-OCT-1999	990S-0160767.			
PR 21-OCT-1999	990S-0160768.			
PR 21-OCT-1999	990S-0160770.			
PR 21-OCT-1999	990S-0160814.			
PR 21-OCT-1999	990S-0160815.			
PR 22-OCT-1999	990S-0160980.			
PR 22-OCT-1999	990S-0160981.			
PR 22-OCT-1999	990S-0160989.			
PR 25-OCT-1999	990S-0161404.			
PR 25-OCT-1999	990S-0161405.			
PR 25-OCT-1999	990S-0161406.			
PR 26-OCT-1999	990S-0161359.			
PR 26-OCT-1999	990S-0161360.			
PR 26-OCT-1999	990S-0161361.			
PR 28-OCT-1999	990S-0161920.			
PR 28-OCT-1999	990S-0161992.			
PR 28-OCT-1999	990S-0161993.			
PR 29-OCT-1999	990S-0162142.			
QY 799	GAGCTACTAGACCCACCACCAATGGTTTCTCCGAGCGCAACTGTGTAGGACAAAGCGGGTTC	858		
Db 454	GAGCTTGAAGCAGCTACTAATGCGATTGTGTGAAGAGAATGTGTATCGGAGGAGGAGGTAT	513		
QY 859	GGTATGTCACAAAGGTGTGTGTCCTAGTGGGAAAGAGTTCTGTGAAAGCAGTTGAAA	918		
Db 514	GGGATTGTTTATAGTGGGATTTAAGTGTAGTGAAGCTATCGGCGGTGTA	633		
QY 919	GTGGGAGTGGTCAGGAGGAGGAGGAGTTCAGGACAGAGGTTGAGATCATCAGCAGATT	978		
Db 574	AACAAGGGGTCAAGCTGAAAGAGGATTTAGAGTGAAGCTATCGGCGGTGTA	633		
QY 979	CACCAAGGCATCTGTGTCTCTGTGTGTGTATTCATCGCCCTGTCACAAAGATTGCTT	1038		
Db 634	AGGCAAGAAGTCTCTAGGCTATTAGGATCTGCGTGAAGGAGCATACAGGATGCTC	693		
QY 1039	GTCATAGTGTTCCTACACACATCTCGAGCTTCACTCCATGCGAGGAGCGGCT	1098		
Db 694	GTGTATGACTAGTGTGATATGGCACTTGGAGCAATGGATTCATGGGATGTGGTAT	753		
QY 1099	ACAA-----TGAATGGAGCACCAGATGAAGATTGCTCTTGGATCTGCTAAGGACTT	1152		
Db 754	AAAGTCCGTTGACTTGGGATATTCGTATGAACATAATACATTGTCATGGCGAAGATTG	813		

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141387.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 05-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      10.2%; Score 199; DB 21; Length 1626;
Best Local Similarity 56.9%; Pred. No. 1.6e-37;
Matches 408; Conservative 0; Mismatches 300; Indels 9; Gaps 2;

QY 759 GAGCTAGCTAGAGCACCAAAAGTTTCTCCAGGCGAACTGTGTAGGACAGCGGGTTC 858
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 454 GAGCTGAGGAGCTACTTAAAGGATTGTGTAAGAGAAATGTATCGGAGAGGAGGTTAT 513
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 859 GGTACGTGCAAAAGGTGTCTGCTAGTGTGGAAGAAGTGTCTGTAAGCAGTTGAAA 918
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 514 GGGATTGTTTATAGTGGGATTTAACTGATGGAATAAGTCGCTGTCAGAAATCTGCTT 573
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 919 GTTGGAGTGTCTAGGAGGAGGAGTTTCAGGACAGGTTGAGATCATCAGCAGAGTT 978
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 574 AACACAGGGGTCTAGCTGAGAGAGGTTTACAGTGGAAAGTGGAAAGCTATCGGGCTGTA 633
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 979 CACCACAGGCATCTGGTCTCTTGTGTTTATTCATCCCGGTCGCCAARAGATGCTT 1038
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 634 AGGATAGAAATCTGTTAGGCTATTAGGATCTGCGTGAAGAGGATCAGAGGATGCTC 693
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1039 GTCTATGAGTTTGTCTTAAACAACATCTCGAGTTTCACCTCCATGCGGAGGAGGCT 1098
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 GTGTATGACTACGTCGATTAATGCGAATCTTGAGCAATGGATTCATGGGATGTTGCTAT 753
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1099 ACAG-----TGGATGAGGACACAGATGTAAGATGCTCTGTGATCTGCTAAAGGACTT 1152
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 754 AAAAGTCGGTTSACTTGGGATATTCGTAAGAACAATAATCTTGCATGCGGAAGGATG 813
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1153 TCTTATCTTCAAGAGATGCAATCTTAATCATCTACCGTGATATCAAGGCTTCAAC 1212
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 814 GCCTATCTGCAAGAGGCTTGAACCGAAGATCTGTCATCGGACATAAATCTAGTAAT 873
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1213 ATATTAATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGCTTGTCTAAGATGCT 1272
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 874 ATTTTCTTGATGCTCAATGGAATGCTAAGGATCAAGATTTGACATGCTTACCTCTTG 933
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1273 TCTGALACAAACAGCATGATATCAACAGCTGTGTGAGGAACTTTGGTACTTGCTCCG 1332
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 TCTCCGAGAGCAGCTACGTACACACCGGAGTAAATGGAATCTTTGGATATGTTGGCT 993
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1333 GAATACGCTGCAAGCGGAAGCTCACGAGAGAGTCTGACGTTTCTCATTTGGCGTTGTG 1392
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 GAGTATGCTTGCACCGGAATGTTGACAGAAAGAGTATATCTATAGTTTGGGATATTA 1053
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 CTTTGGAGCTATTAATGAGCTGACCGCTGATGTCACCAATGCTATGTAGATGAC 1452
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 ATCATGAGATATATCACCGGAAGAAATCTCTGCAATATAGTCGACCTCAAGAGAGTG 1113
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 AGCTTACTTACTGGGACAGCACTTCTT---ACCGAGGATCTGACAGGAGAGAC 1506
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1114 AATTGTTGAGTGGTGAACAAACATGTTGGAAACCGGAGGCTGCTGGAAGTAGTAC 1170
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
AAC45877
ID AAC45877 standard; DNA; 1455 BP.
XX
AC AAC45877;
XX
DT 18-OCT-2000 {first entry}
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48093.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
```

```
KW
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
```

protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.
Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

99US-0121825.

99US-0123180.

99US-0123548.

99US-0125788.

99US-0126264.

99US-0126785.

99US-0127462.

99US-0128234.

99US-0128714.

99US-0129845.

99US-0130077.

99US-0130449.

99US-0130510.

99US-0130891.

99US-0132048.

99US-0132407.

99US-0132484.

99US-0132485.

99US-0132486.

99US-0132487.

99US-0132863.

99US-0134256.

99US-0134218.

99US-0134219.

99US-0134221.

99US-0134370.

99US-0134768.

99US-0134941.

99US-0135124.

99US-0135353.

99US-0135629.

99US-0136021.

99US-0136392.

99US-0136782.

99US-0137222.

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

```
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142377.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153078.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 197.2; DB 21; Length 1455;
Best Local Similarity 57.7%; Pred. No. 4.1e-37;
Matches 374; Conservative 0; Mismatches 268; Indels 6; Gaps 1;

QY 787 TTCATACAGGAGCTAGCTAGAGCCACCACCAATGGTTCTCCGAGGCGCAACTTGTAGGA 846
Db 499 TTACTTTGAGATCTTGAACACAGCACTAACAGATTCTCTAAGAGAATGTATAGGT 558
QY 847 CAAGCGGGTTCGGTTACGTGCACAAAGTGTTGCTAGTGGGAAAGAGTGTCTGTG 906
Db 559 GAAGCGGTTACGGTGTCTACAGAGGAGAGTATGAATGAAGTCTCTGTGTGCAGTT 618
QY 907 AAGCAGTTGAAAGTTGGGAGTGGTTCAGGAGAGAGAGGAGTTTCAGGAGAGGTTGAGATC 966
Db 619 AAAAAGATCTTAACCAAGTGGGACAGCTGAGAAAGAGTTTCAGTCGAGGTTGATGCT 678
QY 967 ATCAGCAGAGTTCCACACAGCACTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1026
Db 679 ATCGGTGATGTGGGCGACAGAACTTGGTTCGGCTTCTTGTGTACTGCAATTGAAGGCACA 738
QY 1027 AAAAGATTGCTGTCTATAGTTGTTCCTAACCAACAATTCGAGCTTCACCTCCATGGC 1086
Db 739 CACAGGATATGGTTTACGATATGTGAATAATGAAGTACATAGACATGCTTCATGGA 798
QY 1087 G-----AGGACGGCCTACAAATGGAATGGAGCACCAGATTGAAGATTGCTTTGGATCT 1140
Db 799 GCAATGAGACAACATGATATCTAACTTGGGAACCAAGGATGAAGTTCTGATCGGCACA 858
```

```
QY 1141 GCTAAGGACTTCTTATCTTCATGAGAGATTGCAATCCTAAATCATTCACCGTGATATC 1200
Db      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
858 TCAAAGCTCTGTCATATCTGTCATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAATG 918
QY 1201 AAGGCTTCAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
919 AAGTCAAGCAATCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 978
QY 1261 GCTAAGATTGCTTCTGATACAAACACGATGATGATGATGATGATGATGATGATG 1320
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
979 GCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
QY 1321 TACTTGGCTCGGTAATACGCTGCAAGCGGAAAGCTCAGGAGAGTCTGACGTTTCTCA 1380
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1039 TACCTGCTCTGTAATGCAAACTCAGGGTTGTTAAATGAAGAGTGTGATGCTACAGC 1098
QY 1381 TTGCGCTGTGCTTTGGAGCTCATTTACTGGAGTGTGACCGCTTGAT 1428
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1099 TTGGTGTGCTTCTTAGAAGCAATCACAGGAGAGACCGCTGTGGAT 1146

RESULT 8
AAC45313
ID AAC45313 standard; DNA; 1751 bp.
XX
AC AAC45313;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46059.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
```

```
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 197.2; DB 21; Length 1751;
Best Local Similarity 57.7%; Pred. No. 4.4e-37;
Matches 374; Conservative 0; Mismatches 268; Indels 6; Gaps 1;

QY 787 TTCACATACAGGAGCTAGCTAGACGACCAACAAATGGTTCTCCGAGGCGAACTTCTTAGGA 846
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 847 CAAAGCGGGTTCGGTTACGCTGCACAAAGGTGTGTGCTAGTGGTAAAGAGTGTCTGTG 906
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 907 AACCAAGTGAAGTGGAGTGTGTCAGGGAGAGAGGAGTTTCAGGCAGAGTTGATC 966
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 967 ATCAGCAGAGTTTCAACACAGGCATCTGGTCTCTTCTTGTGTTTATTCATCGCGGTGC 1026
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 AAAAGATTGCTGTCTATGAGTTTGTCTTAAACAACAACTCGAGCTTCACCTCCATGGC 1086
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1087 G-----AGGGACGGCCTACAATGGTATGAGCAGCACCAGATTGAGATTGCTCTTGGATCT 1140
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 GCTAAAGGACTTCTTATCTTTCATGAGATTGCAATCCTAAAATCATTCACCTGATATC 1200
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1201 AAGGCTTCAACATATGATAGATTTCAGTTTGAAGCTAAGTTGCTGATTTTGGTCTT 1260
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1261 GCTAAGATTGCTTCTGATACAAACACACATGATACACAGTGTGATGGGAACCTTGGG 1320
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1321 TACTTGGCTCCGGATACGCTCAAGCGAAAGCTACGAGAGTCTGACGTTTCTCA 1380
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1381 TTTGGCGTTGCTTTTGGAGCTCATTTACTGGACGTCGACCCGTTGAT 1428
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1428 TTTGGTGTGTTCTCTTAGAAGCAATCACAGGACGACGACCCCTGTGGAT 1207
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9

AAC42469

ID AAC42469 standard; DNA; 1569 BP.

XX AAC42469;

AC AAC42469;

XX AAC42469;

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35678.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130443.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139450.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141342.
PR 02-JUL-1999; 99US-0142154.
PR 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151068.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144874.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 19-MAY-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-015753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.


```
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147535.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-015753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161353.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.0%; Score 194.6; DB 21; Length 1674;
Best Local Similarity 55.9%; Pred. No. 1.8e-36;
Matches 415; Conservative 0; Mismatches 319; Indels 9; Gaps 2;

QY 782 GCACCTTCACATACGAGGAGCTAGCTAGAGCCACCAATGTTTCCTCAGGCGCACTTGT 841
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 553 GGAGCTTCACGTTCAAGGAGTTAGCTGCAGCTACAAGAACTCCGGGAAGTTAATTGC 612
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 842 TAGCAACAGCGGGTTCGGTTACGTCACACAGCTGTGTTCCCTAGTGGGAAGAGTTG 901
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 613 TCGAGGAAGGAGGTTTGGCAGAGTTTATAGGGACGCTTTAGATTACGACAAAGTAGTG 672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 902 CTGTGAAGCAGTTGAAAGTTGGGAGTGTGTCAGGAGAGAGGAGTTTCAGGCAGAGGTTG 961
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 673 CTATTAGCAATTCGAATCCAGATGGGCTTCAAGGAACCGAGAGTTATAGTAGAGTTC 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 962 AGATCATCAGCAGAGTTCCACACAGGCACTGTGTCTCTGTGTTGGTATTGATCGCGG 1021
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 733 TTATGCTTAGCTTATTGCAATCCCAATCTCGTTACATTGATCGGTACTGTACTTCTG 792
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1022 GTGCCAAAGATTCCTTGCTATGAGTTTGTTCCTACACAACTCTCGAGCTTCACTT 1079
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 793 GTGATCAAGAGCTTCTTGCTATGATACATGCAATGGGAAGCTTAGAAGATCACCTTT 852
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1080 ----CCATGGCGAGGAGCGGCTACAATGGAATGGACCAACAGATTGAAGATTGCTCTTG 1135
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


RESULT 13
AAAF77097
ID AAF77097 standard; DNA; 2820 BP.
XX
XX
AC AAF77097;
XX
DT 17-MAY-2001 (first entry)
XX
XX Arabidopsis gene #3.
DE
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance;
KW plant; ds.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO200114563-A1.
PN
XX
XX 01-MAR-2001.
PD
XX
XX 18-AUG-2000; 2000WO-CA00966.
PF
XX
XX 19-AUG-1999; 99US-0149466.
PR
XX
XX 13-OCT-1999; 99US-0159122.
PR
XX
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
XX
XX Goring D, Silva N;
PI
XX
XX WPI; 2001-244305/25.
DR
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen
PT resistance, or for producing transgenic plants with increased wounding
PT or pathogen resistance
XX
XX Examples; Fig 13; 9lpp; English.
PS
XX
XX The present invention relates to proline-rich extensin-like
CC receptor kinase (PERK). The PERK nucleic acids and polypeptides
CC are useful for increasing the resistance of plants to wounding
CC and pathogens. These are also useful for producing transgenic
CC plants with increased wounding and pathogen resistance compared
CC with a wild type plant, as well as in assays for identifying
CC and developing compounds to inhibit and/or enhance polypeptide
CC function directly.
XX
XX Sequence 2820 BP; 819 A; 629 C; 577 G; 795 T; 0 other;
SQ
Query Match 9.9%; Score 192.2; DB 22; Length 2820;
Best Local Similarity 67.2%; Pred. No. 8.2e-36;
Matches 272; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 681 TCACACACCGCTTTCATGAGCAGCAGCGCGCTCCGACTACTCGACCGTCCAGTTCT 740
DB 977 TCACAGACACCGCTTTCATGAGCAGCAGCGCGCTCCGACTACTCGACCGTCCAGTTCT 740
QY 741 TCCTCCACCGCTTTCATGAGCAGCAGCGCGCTCCGACTACTCGACCGTCCAGTTCT 800
DB 1037 TATACCGTCACTCAAGTGCACACTCTTGTCACACACACACACACTTTCACATACGATGA 1096
QY 831 GCTAGTAGACGACCAATGGTTTTCGAGGCGGACCTGTAGGACAGCGGGTTCGG 860
DB 1097 ACTGTCATGCAACAGAGAGTTTCGCTCAGTCAAAATTTGCTAGGACAGAGGATTTGG 1156
QY 861 TTACGTGCAAGAGGTGTGCTAGTGGGAAGAGTTGCTGACAGAGTTGAAGT 920
DB 1157 STATGTTCAAGAGAGTTTCTGCTAGTGGCAAGAGTTGCTGAGAGAGTTTAACT 1216
QY 921 TGGAGTGTGTAGAGAGAGAGAGTTTCAGGACAGAGTTGATCATCAGCAGAGTTCA 980
DB 1217 TGGAGTGTGACAGAGAGAGAGTTTCAGGACAGAGTTGATCATCAGCAGAGTTCA 1276

QY 981 CCACAGGATCGTGTCTCTCTGTTGGTTATTGTCATCGCCGCTGCCAAGAGATTGCTTCT 1040
DB 1277 TCATCGTCATCGCTCGTTTCTCTTGTGGATATTGTCATCTGTTGTCGAAGACTTTTGT 1336
QY 1041 CTATGAGTTTGTCTTCAACAACTCGAGCTTCACCTCCATGG 1085
DB 1337 TTATGAGTTTATACCTACACACACTCTTGAATTCATCTCATGG 1381
RESULT 14
AAAF77096
ID AAF77096 standard; DNA; 2520 BP.
XX
XX AAF77096;
XX
DT 17-MAY-2001 (first entry)
XX
XX Arabidopsis gene #2.
DE
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance;
KW plant; ds.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO200114563-A1.
PN
XX
XX 01-MAR-2001.
PD
XX
XX 18-AUG-2000; 2000WO-CA00966.
PF
XX
XX 19-AUG-1999; 99US-0149466.
PR
XX
XX 13-OCT-1999; 99US-0159122.
PR
XX
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
XX
XX Goring D, Silva N;
PI
XX
XX WPI; 2001-244305/25.
DR
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen
PT resistance, or for producing transgenic plants with increased wounding
PT or pathogen resistance
XX
XX Examples; Fig 12; 9lpp; English.
PS
XX
XX The present invention relates to proline-rich extensin-like
CC receptor kinase (PERK). The PERK nucleic acids and polypeptides
CC are useful for increasing the resistance of plants to wounding
CC and pathogens. These are also useful for producing transgenic
CC plants with increased wounding and pathogen resistance compared
CC with a wild type plant, as well as in assays for identifying
CC and developing compounds to inhibit and/or enhance polypeptide
CC function directly.
XX
XX Sequence 2520 BP; 717 A; 553 C; 537 G; 713 T; 0 other;
SQ
Query Match 9.9%; Score 192; DB 22; Length 2520;
Best Local Similarity 70.8%; Pred. No. 8.8e-36;
Matches 255; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 730 CGTCCAGTTCTTCTCCACCGTCTCCAGGCTTGTGTAGGCTTCTCCAAAGCATTTC 789
DB 841 CGCCCACTTACTCTCTCTCCCTGCTCTAGCCCTCGATTCACACAGACACTTT 900
QY 790 ACATACAGGAGCTAGTAGACCCACCAATGGTTTCTCCGAGGCGAATTTGTAGGACAA 849
DB 901 ACTTACCAAGAGCTTGGGCTGCAACAGAGGTTTACGGATGCTAACCTTTTGGGACAG 960
QY 850 GCGGGTTCGTTACGTGCAACAGGTTTCCCTAGTGGGAAGAGTTGCTGCTGAG 909

```
Db 961 GGAGGATTTGGTATGTCATCAAGGAGTCTTGCCTAGCGGGAAGAGTAGCAGTTAAG 1020
QY 910 CAGTTGAAAGTTGGGAGTGGTCAGGAGAGAGAGGAGTTTCAGGCAGAGTTGAGATCATC 969
Db 1021 AGTTTAAAGCGGTAGCGCAAGAGAGAGAGGAGTTTCAGCTGAGTCCGATATTCATT 1080
QY 970 ACAGAGTTCCACACAGGAGCATCTGGTGTCCTCTGTTGTTGTTATGATCCCGTGCCAAA 1029
Db 1081 ACCCGTGTGCATCATCGGTATCTGTTCTTTGGTTGGATATGTCATGATGGACAG 1140
QY 1030 AGATTGCTGTCTATGAGTTTCTTCTACACAAATCICGAGCTTCACCTCCATGGCGAG 1089
Db 1141 AGAGTTGGTTATGAGTTTCTTCTACAAACATTTGGATATCATCTTCATGTTAG 1200

RESULT 15
AAC50527
ID AAC50527 standard; DNA; 1848 BP.
XX AAC50527;
AC
XX
XX
DF 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 65175.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; terminator; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134220.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
```


D _b	121	CGATCCACATATTCGGAGATCTCCTCCCTTCTTCTGCTACACCTTTGCCTCTCTCT	180
Q _y	181	CCATCTCCACCAACTCCATCTACGCCGGGATCTCCACCTCTCTCTTCTCACCGGTCTCCA	240
D _b	181	CCATCTCCACCAACTCCATCTACGCCGGGATCTCCACCTCTCTTCTCACCGGTCTCCA	240
Q _y	241	CCCGTCTCAACTAGCGCCGGATCTCCACCGCACTGTATTACTCTCTTACTCTCGAAACGCT	300
D _b	241	CCCGTCTCAACTAGCGCCGGATCTCCACCGCACTGTATTACTCTCTTACTCTCGAAACGCT	300
Q _y	301	CCACCTTCAGTCCOAGGACCACGTCOAATCTTACGCGAAGAGAGATCTCCCTCGACCT	360
D _b	301	CCACCTTCAGTCCOAGGACCACGTCOAATCTTACGCGAAGAGAGATCTCCCTCGACCT	360
Q _y	361	CCATCTTCTCCCTCGCGCGCTCTCTCTTCTTCOGACGGTTTATCAACAGGAGTGGTGGTG	420
D _b	361	CCATCTTCTCCCTCGCGCGCTCTCTCTTCTTCOGACGGTTTATCAACAGGAGTGGTGGTG	420
Q _y	421	GGAATCGCCATCGSAGGAGTCGCTCTGCTTGTGATAGTAGCTCTGATTTGTCTCTCTGT	480
D _b	421	GGAATCGCCATCGSAGGAGTCGCTCTGCTTGTGATAGTAGCTCTGATTTGTCTCTCTGT	480
Q _y	481	AAGAAGAAACGACGGAGAGACGAAGAAGATGCTTACTATGTTCTCTCCGCACTCTCTCT	540
D _b	481	AAGAAGAAACGACGGAGAGACGAAGAAGATGCTTACTATGTTCTCTCCGCACTCTCTCT	540
Q _y	541	GGTCCGAAAACGCGGAGGACCTTACGGTGGACAGCAGCAACAATGGCGGCAACAAACGCA	600
D _b	541	GGTCCGAAAACGCGGAGGACCTTACGGTGGACAGCAGCAACAATGGCGGCAACAAACGCA	600
Q _y	601	ACACCAACCGTCAGATCATGTCGTGAGTCCTACTACCAACCACTAAGGCTCCATCTCCA	660
D _b	601	ACACCAACCGTCAGATCATGTCGTGAGTCCTACTACCAACCACTAAGGCTCCATCTCCA	660
Q _y	661	CCACGCAACCTCTCTCACTCCACCAACCGCTTCTATGACGACGACGCGCGGCTCCGAC	720
D _b	661	CCACGCAACCTCTCTCACTCCACCAACCGCTTCTATGACGACGACGCGCGGCTCCGAC	720
Q _y	721	TACTCGAACCGTCCAGTCTCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA	780
D _b	721	TACTCGAACCGTCCAGTCTCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA	780
Q _y	781	AGACATTCACATACAGGAGCTAGCTAGAGCCACCAATGTTTTCTCCGAGGCGAAGTTG	840
D _b	781	AGACATTCACATACAGGAGCTAGCTAGAGCCACCAATGTTTTCTCCGAGGCGAAGTTG	840
Q _y	841	TTAGGCAAGCGCGGTTCCGTTACGTCSCAAAAAGGTGTGTGCTAGTGGAAAGAGTT	900
D _b	841	TTAGGCAAGCGCGGTTCCGTTACGTCSCAAAAAGGTGTGTGCTAGTGGAAAGAGTT	900
Q _y	901	GCTGTGAACAGTTGAAGTTGGAGTGGTTCAGGGAGAGAGGAGTTTCAGGACAGGTT	960
D _b	901	GCTGTGAACAGTTGAAGTTGGAGTGGTTCAGGGAGAGAGGAGTTTCAGGACAGGTT	960
Q _y	961	GAGATCATCAGCAGAGTCTACCAACAGCANCTGGTGCTCTGTTGTGTTATTCGATCGCC	1020
D _b	961	GAGATCATCAGCAGAGTCTACCAACAGSCANCTGGTGCTCTGTTGTGTTATTCGATCGCC	1020
Q _y	1021	GGTGCCAAAAGATPGCTTGTCTATGASTTTGTTCTTAACAACAATCTCGAGCTTCACTIC	1080
D _b	1021	GGTGCCAAAAGATPGCTTGTCTATGASTTTGTTCTTAACAACAATCTCGAGCTTCACTIC	1080
Q _y	1081	CATGGCAGGCGCGCTACAAATGGAAATGGAGCACCAGATTTGAAGATGCTCTTGNATCT	1140
D _b	1081	CATGGCAGGCGCGCTACAAATGGAAATGGAGCACCAGATTTGAAGATGCTCTTGNATCT	1140
Q _y	1141	GCTAAAGGACTTTCTTAICTTCATGAAGATTGAATCTCAATCTAAATCATCCCGTGATC	1200
D _b	1141	GCTAAAGGACTTTCTTAICTTCATGAAGATTGAATCTCAATCTAAATCATCCCGTGATC	1200
Q _y	1201	AGGCGTTCAACANATTTGATGAGATTTCAAAGTTTGAAGTTGCTGATTTTGGTCTT	1260

Db	1201	AAGCTTTCACACATATTGATAGATTTCAAGTTTAAAGTTGCTGATTGGTCTT	1260
QY	1261	GCTAAGATTGCTTCTGATACAAACACGCATGTATCAACACGCTGTGATGGAACTTTGGG	1320
Db	1261	GCTAAGATTGCTTCTGATACAAACACGCATGTATCAACACGCTGTGATGGAACTTTGGG	1320
QY	1321	TACTTGGCTCCGGAATACCGTGCACGCGGAAGCTCACGGAAGAAGTCTACGCTTTTCTCA	1380
Db	1321	TACTTGGCTCCGGAATACCGTGCACGCGGAAGCTCACGGAAGAAGTCTACGCTTTTCTCA	1380
QY	1381	TTTGGGCTTGCTTTTGGAGCTCATTTACTGGAGCTGCACCGTTGATGCCAACAAATGTC	1440
Db	1381	TTTGGGCTTGCTTTTGGAGCTCATTTACTGGAGCTGCACCGTTGATGCCAACAAATGTC	1440
QY	1441	TATCTAGATGACAGCTTAGTTGACTGGGACGACCAATTCCTTAAACGGACATCTGAGCA	1500
Db	1441	TATCTAGATGACAGCTTAGTTGACTGGGACGACCAATTCCTTAAACGGACATCTGAGCA	1500
QY	1501	GGAGACTTTGAGGGTTTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGAGATG	1560
Db	1501	GGAGACTTTGAGGGTTTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGAGATG	1560
QY	1561	GCTCGCATGGTTGCTTGTGCTGGGGCTTGTGTTTGGCCATTACGTCGCCGAGACCTGCG	1620
Db	1561	GCTCGCATGGTTGCTTGTGCTGGGGCTTGTGTTTGGCCATTACGTCGCCGAGACCTGCG	1620
QY	1621	ATGAGCCAGATTCTGGTGGCTTAGAAGAGAAATGTTATCACTCTCAGATCTTAAACGAAGG	1680
Db	1621	ATGAGCCAGATTCTGGTGGCTTAGAAGAGAAATGTTATCACTCTCAGATCTTAAACGAAGG	1680
QY	1681	ATGAGACCAGGTCAAAGCAATGTATACGTCCTATCGGAGGAAGCACCGATTATCACTCG	1740
Db	1681	ATGAGACCAGGTCAAAGCAATGTATACGTCCTATCGGAGGAAGCACCGATTATCACTCG	1740
QY	1741	AGCCATGACANTGAAGACATGAAGAAGTTTAGGAAAATGGGACTTTGGAACTCAAGAGTAC	1800
Db	1741	AGCCATGACANTGAAGACATGAAGAAGTTTAGGAAAATGGGACTTTGGAACTCAAGAGTAC	1800
QY	1801	ACGCCACGGGTGAGTACAGTAAATCCGACCAGTGTATGSHACTGTACCCGCTCGGTTCA	1860
Db	1801	ACGCCACGGGTGAGTACAGTAAATCCGACCAGTGTATGSHACTGTACCCGCTCGGTTCA	1860
QY	1861	AGCAGCGAGGGCCAAACACACCGGAAATGGAGATGGGGAAGATTAGAGAACCGGTCAG	1920
Db	1861	AGCAGCGAGGGCCAAACACACCGGAAATGGAGATGGGGAAGATTAGAGAACCGGTCAG	1920
QY	1921	GGTTATAGTGGACCTTCTTTAA	1944
Db	1921	GGTTATAGTGGACCTTCTTTAA	1944

```

; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2189)
; OTHER INFORMATION:
DS-10-086-454-3

```

Query Match	100.0%;	Score 1944;	DB 9;	Length 2189;
-------------	---------	-------------	-------	--------------

[illegible]

```
US-10-086-464-10
; Sequence 10, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 10
; LENGTH: 1902
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1902)
US-10-086-464-10

Query Match      26.6%; Score 516.2; DB 9; Length 1902;
Best Local Similarity 67.1%; Pred. No. 1.5e-139;
Matches 747; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

QY 730 CGTCCAGTCTTCTCCACCGCTCCAGGGCTTGTAGGCTTCTCCAAAAGACACTTTC 789
DB 757 CCGCCAGTCTTCTCCACCGCTCCAGGGCTTGTAGGCTTCTCCAAAAGACACTTTC 816

QY 790 ACATACAGAGAGTAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
DB 817 ACTTACCAAGAGGTGCGGCTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 876

QY 850 GCGGGTTCGGTACGTCGCAAGAGTGTCTTCCTAGTGGGAAAGAGTGTCTGTCGAAG 909
DB 877 GGAGGATTTGGGTATGTCATTAAGGAGTCTTCCCTAGCGGGAAGAGTAGCGAGTAA 936

QY 910 CAGTTCAGAGTTCGGAGTGTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969
DB 937 AGTTTAAAGCGGGTAGCGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 996

QY 970 AGCAGAGTTCACACAGGAGTCTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1029
DB 997 AGCGGTGTGATCATCGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1056

QY 1030 AGATTCGTCTATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
DB 1057 AGGATCTGTGATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1116

QY 1090 GGACGGCTACATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1149
DB 1117 AATCTTCGGTAATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1176

QY 1150 CTCTCTATCTCTATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1209
DB 1177 CTGCTTACCTTCAAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1236

QY 1210 AACATATGATAGATTTCAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGA 1269
DB 1237 AATATTCITGACTTCACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1296

QY 1270 GCTTCATACAAACAGGAGTATCAACAGGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGT 1329
DB 1297 ACATCTGATAACAACTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1356

QY 1330 CCGGATACGCTGCAAGCGGAAGAGTCTACGAGAGTCTACGAGTCTCTCTCTCTCTCTCTCT 1389

US-10-086-464-9
; Sequence 9, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-086-464-9

Query Match      26.6%; Score 516.2; DB 9; Length 1939;
Best Local Similarity 67.1%; Pred. No. 1.5e-139;
Matches 747; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

QY 730 CGTCCAGTCTTCTCCACCGCTCCAGGGCTTGTAGGCTTCTCCAAAAGACACTTTC 789
DB 774 CGCCAGTCTTCTCCACCGCTCCAGGGCTTGTAGGCTTCTCCAAAAGACACTTTC 833

QY 790 ACATACAGAGGAGTAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
DB 834 ACTTACCAAGAGTGTGGCTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 893
```


Db 1791 ACAGCACTAGAGGAGATATGTCATGATGATCTAAGTGAGGAGCAACAGACAGGACA 1850
 QY 1695 AAGCAATGTATACAGCTCATACGAGGAGAGCAGCAGGATATGACTCGAGCCAGTACATGA 1754
 Db 1851 AAGCACAGTACTGAGCCCGGGAGCGGTGAGCTCAGAGTATGACCAAGCTCGTACACGGC 1910
 QY 1755 AGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTACACGCGCAGGSEGA 1814
 Db 1911 AGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTACACGCGCAGGSEGA 1966
 QY 1815 GTACAGTAATCCGACCACTGACTATGAGTACTGACCGCTGTACCGCTGGTTCACAGCAGGAGGSCCA 1874
 Db 1967 --AATGTTGGTGAACAAGTGAATGATGCTTAAACCCCTTCGCTTCAAGTAGTGAAGAAT 2024
 QY 1875 AACACACAGCGCAATGGA 1892
 Db 2025 GAATAGAGGCTCAATGAA 2042

RESULT 7

US-10-086-464-16
 ; Sequence 16, Application US/10086464
 ; Publication No. US20020199218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GORING, Daphne R. et al.
 ; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
 ; FILE REFERENCE: P 25,762-A USA
 ; CURRENT APPLICATION NUMBER: US/10/086,464
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 10/069,304
 ; PRIOR FILING DATE: 2002-02-19
 ; PRIOR APPLICATION NUMBER: PCT/CA00/00966
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: US 60/149,466
 ; PRIOR FILING DATE: 1999-08-19
 ; PRIOR APPLICATION NUMBER: US 60/159,122
 ; PRIOR FILING DATE: 1999-10-13
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 16
 ; LENGTH: 2196
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2196)
 US-10-086-464-16

Query Match 21.3%; Score 413.6; DB 9; Length 2196;
 Best Local Similarity 67.0%; Pred. No. 1.2e-109;
 Matches 606; Conservative 0; Mismatches 289; Indels 9; Gaps 1;
 QY 782 GCACCTTCATACAGGAGGAGCTAGCTAGAGCCACCAATGGTTCTCCGAGGCGAAGCTGT 841
 Db 1124 GAATGTTCTCTCAGAGAACTTCCAAAAGCACTGGTGGATTTTCAGAGGAGAACCTTT 1183
 QY 842 TAGCACAAGCGGGTTTCGGTACCTGCACAAAGTGTGTTCCTAGTGGGAAAGCTTG 901
 Db 1184 TGGGAAAGCGGGTTTCGGATATGTTCCAAAAGGAGTGTTCAAAAGCGGGACAGAGTTG 1243
 QY 902 CTGTGAAGCAGTGTGAAGTTGGGAGTGGTCAGGAGGAGAGGAGCTTTCAGGACAGGTTG 961
 Db 1244 CGGTGAAGCAGTGTGAAGTTGGGAGTGTTCAGGAGGAGAGAGATTCAGCTAGGTTG 1303
 QY 962 AGATCATACAGAGTTTACACAGGAGCATCTGGTCTCTCTTGTGGTTATTCATCGATCCG 1021
 Db 1304 ACACAACTAGTAGGTTTCATCAATGAACACCTCGTTTCATTGTTGGTTATTGCGTTAATG 1363
 QY 1022 GTGCCAAAAGATTCCTGTCTATGAGTTTGTTCACACAACTCTCGAGCTTCACCTCC 1081
 Db 1364 GAGATTAAGAGACTCTTGTGTTACAGTTTGTTCCTTAAGATACCTTCGGAGTTCACCTGC 1423
 QY 1082 ATGGCGAGGAGCGGCTTCAATGAATGAGTGGAGCAGACCAAGATTGAAGATTGCTCTTGATCTG 1441

Db 1424 ATGAGAAACAGAGGAGCGGTGTTGGATGGGAATGAGGCTCAGGATGCTGTAGGAGCAG 1483
 QY 1142 CTAAGGAGACTTCTCTATCTTCATGAGATTCGAAATCTCTAAATCATCTACCGTGTATCA 1201
 Db 1484 CAAAAGGATTAGCTTATCTTCATGAGGATTCGAGTCCAACTATAATTCACCGTGTATCA 1543
 QY 1202 AGGCTTCAACATATTGATAGATTTCAGTTTTCAGCTAAGCTTCTGATTTTGGTCTTG 1261
 Db 1544 AAGCAGCTATATCTCTTAGATTCCAAATTTGAGGCAAGGCTCTCTGACTTTGGACTAG 1603
 QY 1262 CTAAGATTGCTTCTCATACAAA-----CAGCAGTATATCAACACCTGTGTATGSGAA 1312
 Db 1604 CCAAGTTTTCACACACCAATTCATTCATCTACTCATACTCTACTCAGTGGTAGSAA 1663
 QY 1313 CTTTGGGTACTTGGCTCCGGAATAGCTGCAAGCGGAAAGCTCAGGAGAGAGTCTGACG 1372
 Db 1664 CTTTGGGATACATGGCTCCAGAAATACGGCTCCAGTGGTAAAGTAACTGATAAATCAGATG 1723
 QY 1373 TTCTCTCATTTGGCGTTGCTTTTGGAGCTCAITACTGAGGCTCGACCCGTTGATGCCA 1432
 Db 1724 TATATTCCTTTGGGCTCGTCTTAGAACTCATCTAGGAGCTCCATCAATTTGCGCCA 1783
 QY 1433 ACAATGCTCTATGATGACAGCTTAGTTGACTGGCGACGACCATTCGTTAAACCGAGCAT 1492
 Db 1784 AAGATCTTCCACAACACAGAGTTTAGTAGACTGGCGAGGCCATTCGTTACAGAGCA 1843
 QY 1493 CTGACAAAGGAGACTTTGAGGGTTAGCTGATGATCAAGAAGATGAATATGGGTATGACAGAG 1552
 Db 1844 TCCTCGGAGAAAGTTTTCAGCTTCTTAGACTCAAGGTTGGAGAAGAAATTACGATACAA 1903
 QY 1553 AGGAGATGGCTCGCATGGTTGCTTGTGCTGGCGCTTGTTCGCCATTCAGCTCCGCGCA 1612
 Db 1904 CTCAGATGGCAAAACATGGCTGCTGTGCTGCTGCTTGCATACGCCAATCAGCTTGCCTTC 1963
 QY 1613 GACCTCGCATGAGCCAGATTTGCTGCTGCTTGAAGGAAATATATCACTGTCTCAGATCTTA 1672
 Db 1964 GGCCTAGATGAGCCAGCTAGTACGCTCTCTTGAAGCGAGGTGGCCCTGAGAAAGGTGG 2023
 QY 1673 ACGA 1676
 Db 2024 AAGA 2027

RESULT 8

US-10-086-464-15
 ; Sequence 15, Application US/10086464
 ; Publication No. US20020199218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GORING, Daphne R. et al.
 ; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
 ; FILE REFERENCE: P 25,762-A USA
 ; CURRENT APPLICATION NUMBER: US/10/086,464
 ; CURRENT FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 10/069,304
 ; PRIOR FILING DATE: 2002-02-19
 ; PRIOR APPLICATION NUMBER: PCT/CA00/00966
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: US 60/149,466
 ; PRIOR FILING DATE: 1999-08-19
 ; PRIOR APPLICATION NUMBER: US 60/159,122
 ; PRIOR FILING DATE: 1999-10-13
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 2261
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-086-464-15

Query Match 21.3%; Score 413.6; DB 9; Length 2261;
 Best Local Similarity 67.0%; Pred. No. 1.3e-109;
 Matches 606; Conservative 0; Mismatches 289; Indels 9; Gaps 1;


```
QY 782 GCACTTTACATACGAGGAGTAGAGCCACCAATGTTCTCCGAGCGCACTTGT 841
Db 1151 GAATGTTCTCTCAGAGAACTTCAAAGCACTGGTGGATTTTCAGAGGAGACCTTT 1210
QY 842 TAGGACAAAGGGGTTGGTTACGTGCACAAAGTGTTCGCCAGTGGGAAAGATTG 901
Db 1211 TGGGAGAGGGGTTGGATATCTTCACAAAGGAGTGTTCGAAACCGGACAGAGTTG 1270
QY 902 CTGTGAACCACTTAAAGTTGGAGTGGTCAAGGAGAGGAGGTTTCAGGACAGGTTG 961
Db 1271 CGGTGAACCACTGAAGATTGGGAGCTATCAAGGAGGAGGAGGAGGTTTCAGGTTG 1330
QY 962 AGATCATCAGCAGAGTTCAACACAGGAGTGTCTCTGTTGGTTATTGCATCGCG 1021
Db 1331 ACACAACTAGTAGGTTCAATAGCACCCCTGTTTCATGTTGGTTATTGCGTTAATG 1390
QY 1022 GTGCCAAAGATTGCTGTCTATGAGTTTGTTCACAACTTCAGGCTTCACCTCC 1081
Db 1391 GAGATAAAGACTCTTGTTTACGAGTTGTCTTAAAGATACCTTGGAGTTCCACTGC 1450
QY 1082 ATGCGGAGGAGCGGCTTACATGAGTGAATGAGCAGCAGATTGAAGATGCTTGCATG 1141
Db 1451 AIGAGACAGGAGGAGGCTGTGGAATGGGAAAGAGGCTCAGGATTGCTGTAGGAGCAG 1510
QY 1142 CTAAGAGACTTCTTATCTTCATGAAGATTGCAATCCTTAAATCATTCACCGTGATCA 1201
Db 1511 CAAAAGGATTAGCTTATCTTCATCAGGATTGCAATCCTTAAATCATTCACCGTGATCA 1570
QY 1202 AGGCTTCAACATATTGATGATTTCAAGTTTGAAGTTGAAGTTGCTGATTTGCTGTTG 1261
Db 1571 AAGCAGCTAATATCTCTAGATTCCAAATTTGAGCAAAAGGCTCTGACTTGGACTAG 1630
QY 1262 CTAAGATTGCTCTGATACAAA-----CAGCATGTATCACAGTGATGAGGAA 1312
Db 1631 CCAATTTCTCAGACACCAATTCATCTCACTCATATCTCTACTCGAGTGGTAGGAA 1690
QY 1313 CTTTGGGTACTTGGCTCCGGAATACGCTGGAAGGAAAGCTCACGGAGAGTCTGAGC 1372
Db 1691 CTTTGGGATACATGGCTCCAGATACGCTGCAAGTGGTAAAGTAACTGATTAACAGATG 1750
QY 1373 TTTTCTCATTTGGGTTGTGCTTTTGGAGCTCATCTAGGAGCTGAGCCCTGATGCGCA 1432
Db 1751 TATATCTTCTTGGGCTGCTGCTTCAAGAACTCATCTAGGAGCTGAGCCCTGATGCGCA 1810
QY 1433 ACAATGCTATGTAGATGACAGCTTAGTGTAGCTGGCAGGAGGAGTGTAAACGAGCAT 1492
Db 1811 AAGATTCTTCCACAAACAGAGTTAGTAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1870
QY 1493 CTGAGCAAGAGACTTTGAGGTTTGTAGTGAATGCAAGAGATGAATGAGTATGAGCAG 1552
Db 1871 TCTCTGAGAGAAAGTTTGAATTTCTTGTAGACTCAAGTTTGGAGAGAAATACGATACAA 1930
QY 1553 AGAGATGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1612
Db 1931 CTGATGAGCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1990
QY 1613 GACCTGCTGAGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672
Db 1991 GGCCTAGAAAGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050
QY 1673 ACAG 1676
Db 2051 AAGA 2054
```

RESULT 9

US-10-086-464-7

; Sequence 7, Application US/10086464

; Publication No. US20020199218A1

; GENERAL INFORMATION:

; APPLICANT: GORING, Daphne R. et al.

; TITLE OF INVENTION: PROLINE-RICH EXIENSIN-LIKE RECEPTOR KINASES

```
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2196)
US-10-086-464-7
```

```
Query Match 15.7%; Score 305.6; DB 9; Length 2196;
Best Local Similarity 59.1%; Pred. No. 3.6e-78;
Matches 644; Conservative 0; Mismatches 344; Indels 102; Gaps 3;

QY 786 TTTCATACAGGAGGAGCTAGTAGAGCCACCAATGTTCTCCGAGGAGCACTTGTAGG 845
Db 1074 TTTCAGTTACGAGAGCTTGTGTAGATACACAGGCTTGTCTCGCAAAACATCTTTGG 1133
QY 846 ACAAGCGGGTTCGTGTAGTGCACAAAGTGTGTGCTGAGTGGGAAAGAGTGTCTGT 905
Db 1134 AGAAGCGGATTTGATGTGTCTATAGAGTACATTCAGGATGTAAAGTTGTTCGGT 1193
QY 906 GAAGCAGTTGAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGAGAGGTTGAGAT 965
Db 1194 TAAGCAGCTTAAGCTGGAAGTGGACAAAGTGACCGTGAATTCAAAGCAGAGGTTGAGAT 1253
QY 966 CATCAGAGAGTTCACACAGGAGTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1025
Db 1254 CATCAGCGCGTTCATCATCGCCATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1313
QY 1026 CAAAAGATTGCTTCTATGAGTTTCTTAAACAAATCTCGAGCTTCACCTCCATGG 1085
Db 1314 GCATAGATTTCTATCTATGATGTATGTTCTAATCAACCTTGGAGCAATTCGAT--- 1371
QY 1086 CGAGGAGCGGCTACATGGAATGAGCAGCAGATTTGAAGATTTGCTTTGGATCTGTAA 1145
Db 1372 -----GAGTGTCTAAGAGAGTCCGGATCGCTATAGGATCAGCCAA 1412
QY 1146 AGGACTTCTTATCTTCATGAGATTCGAATCCATCCATTAATCTTCACCGTGTATCAAGGC 1205
Db 1413 AGGTTGGCATCTCTCAGAGAGCTGTCTCCGAAATCATTCACAGAGATATAAGATC 1472
QY 1206 TTCAACATATTGATAGATTTCAAGTTTGAAGCTAAG----- 1242
Db 1473 AGCAATATTCTCTAGTGTGATGATATGAAGCTCAGGCAATTAAGAAATCTCTCTTTC 1532
QY 1243 -----GTTGCTGATTTGGTCTTGTGTAAAGTTGC 1271
Db 1533 GTTAAATCTATCTTATGACTGTAAAGTTTATAGTTTGTGATTTGGACTGTGTAGACTCAA 1592
QY 1272 TTCTGATACAAACAGGATGTATCAACAGCTGTGATGGAGACCTTTGGTGTGTGGTCC 1331
Db 1593 TGATACAAACAACTCATGTTTCACTCGGGTTATGGGAACCTTCGGGTACCTAGCGCC 1652
QY 1332 GGAATAGCTTCAAGCGGAAAGCTCAGGAGAGGCTCTGACCTTTTCTCATTTGGCGTTGT 1391
Db 1653 GGAATATGATCAAGTGGAAATTCAGTGTAGATCCGATGTATCTCATTCATTCGGGGTGT 1712
QY 1392 GCTTTTGGAGCTCAATCTGAGCAGCTGACCGGTTGATGCCAACAAATGTCTTAAGTAGATGA 1451
Db 1713 TCTCTTAGAGCTTGTAACTGACGGAACCAAGTTTGACAGACTTCAGGCTCTAGGAGAAGA 1772
```


RESULT 13

US-09-938-842A-1911
; Sequence 1911, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1911
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1911

Query Match 10.2%; Score 199; DB 9; Length 1353;
Best Local Similarity 56.9%; Pred. No. 3.1e-47;
Matches 408; Conservative 0; Mismatches 300; Indels 9; Gaps 2;

QY 799 GAGTACCTAGAGCCACCACATGTTCTCCGAGGCGACTTGTAGACAAAGCGGGTTC 858
DB 454 GAGCTTGAACACACTACTAATGATGTTGTGAAGAGAAATGATCGGAGAGGAGGTAT 513
QY 859 GGTACCTGACAAAGAGTGTGCTAGTGGGAAGAAGTGTGTAAGACAGTTCGAAA 918
DB 514 GGGATGTTTATAGTGGGATTTAATGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 573
QY 919 GTTGGAGTGGTCAGGAGAGAGGAGTTCAGGAGAGGTTGAGATCATCAGCAGAGTT 978
DB 574 AACACAGAGGCTCAAGCTGAGAGGAGTGTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 633
QY 979 CACCAAGGATCTGGTCTCTCTGTTGGTATGATGCTATGCTGCTGCTGCTGCTGCTGCT 1038
DB 634 AGGCATAAAGATCTCGTGGCTATTAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
QY 1039 GTCATGAGTTCCTTCTTACAAACAATCTCGAGCTTCACCTCCATGCGGAGGACGGCT 1098
DB 694 GTCATGACTACCTCGATATGSCAATCTGGAGCAATGATTCATGGGATGTTGGTAT 753
QY 1099 ACA-----TGAATGGACACACAGATTGAAGATGCTCTTGGATCTGTAAGGACTT 1152
DB 754 AAAAGTCCCTGACTTGGATATTCGTATGAACATAATATCTTGCATGGGGAAGGATTG 813
QY 1153 TCTATCTTCATGACATGCAATCTCTAAATCATTCACCTGCTGATATCAGGCTTCAAC 1212
DB 814 GCCTATCTGATGGGCTCTTGACCCGAAAGTCTGCTTCGCGGACATAAATCTAGTAAT 873
QY 1213 ATATTGATAGATTTCAAGTTTGAAGTAAAGTGTGCTGATTTGGCTTGTGTAAGATGCT 1272
DB 874 ATTTTGCTTGATCGTCAATGGAATGCTAAGTATCAGATTTTGGACTTGTAGCTTTG 933
QY 1273 CTGATACAAACAGGATGATCAACACGTTGTATGGGAACCTTTGGTACTTGGCTCCG 1332
DB 934 TCTCCGAGAGCAGCTACGTGACAAACCGAGTAATGGGAATTTGGATATGTTGGCCT 993
QY 1333 GAATAGCTCCAGCGGAAGCTCACGGAAGTCTGACGCTTTCTCATTTGGCTTGTG 1392
DB 994 GAGTATGCTTGCACCGGAATGTTGACAGAAAGAGTGATATCATAGTTTGGGATATTA 1053
QY 1393 CTTTGGAGCTCATTTACTGGACGCTGACCGCTTGTGATGCCAACAAATGCTATGATGAC 1452

DB 1054 ATCATGGAGATAATCACCGGAAGAAATCTCTGCAATTATAGTCGACCTCAGGAGGTTG 1113
QY 1453 AGCTTAGTTGACTGGCAGCAGCACCATTCTT---AACCGAGCATCTGACCAAGGAGAC 1506
DB 1114 AATTTGGTTGAGTGCTTGAAACAAATGTTGAAACCGAAGTCTCTGAGGAAGTAGTC 1170

RESULT 14
US-09-887-576-179/c
; Sequence 179, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-179

Query Match 10.0%; Score 194.8; DB 10; Length 2004;
Best Local Similarity 55.5%; Pred. No. 6.6e-46;
Matches 401; Conservative 0; Mismatches 312; Indels 9; Gaps 1;

QY 799 GAGCTAGCTAGAGCCACCACATGTTCTCCGAGGCGAACTTGTAGGACAAAGCGGGTTC 858
DB 991 GAGCTTAGGTTCTACTAATGGAATTCCTGACGAGAATGATTTGGACAGAGGAGTTAT 932
QY 859 GGTACGTCGCAAAAGGTGTGTTGCTAGTGGGAAGAAGTGTGCTGTTGGAAGCAGTTGAAA 918
DB 931 GGGATTGTGTATAGAGTGTCTTTGAGGATAAATCAATGGTGGCTATAAAGAAATCTGCTC 872
QY 919 GTTGGAGTGTCTGAGGAGAGGAGGAGTTTCAGGACAGAGGTGAGATCATCAGCAGAGTT 978
DB 871 AACACAGAGACAGCTGAGGAAGAATTTAAAGTCGAAGTCGAAGCTATTGGAGCCGTA 812
QY 979 CACCAAGGCTATGCTGCTCTCTTGTGTTATTCATCGCGGTGCGCAAAAGATTCGTT 1038
DB 811 AGACACAGATCTGGTGAATGCTTGGTTATTCGTTGAGAGGACTCATAGGATGTA 752
QY 1039 GTCATGAGTTGTTTCCCTAACAAATCTCGAGCTTCACCTCCATGCGGAGGG----- 1091
DB 751 GTTTACGAGTATGTTGATAATGGGAATTTAGAACAAATGGAATTCATGGTGGTTGGGG 692
QY 1092 --ACGGCTACAATGGATGGAGCAGCAGATGGAATGGAATTCCTTGGATCTGCTAAAGTA 1149
DB 691 TTTAAGAGTCTCTTACTTGGAGATTAGGATGACATTTGTTTGGACAGCCAAAGGG 632
QY 1150 CTTCCTTATCTCATGAGATGCAATCTTAAATCAATCCCGTATCAAGGCTTCA 1209
DB 631 TTAATGATTATTACAGAGGCTTTGAGCCTTAAGTTGTTCCAGGAGATTAATATCGAGT 572
QY 1210 AACATATTGATAGATTTTCAAGTTTGAAGTAAAGTGTGCTGATTTTGGCTTGTGAAGATT 1269
DB 571 AATATCTGCTTGACAGCAATGGNACTCGAAAGTTCGAAAGTTCGACTTTGGGTAGTAGCTC 512

QY	1270	GCTTTCIGATACAAACACGCATGTATCAACACGCTGTGATGGAACTTTGGTACTTGGCT	1329
Db	511	TTGGGTTGCGAGATGACCTATGTGTCACCTACCTACCTGCTGTAAGGATCGTTTGGGTATGTGGCT	452
QY	1330	CCGGAATACCGTGCACCGGGAAGCTCACGGAGAAGTCTACGCTTTTTCATTTTGGCGTT	1389
Db	451	CCTGAATACCGAGTACCGGAATGTTGAATGAGAGGACGATGTTTATAGCTTTCGGGTGT	392
QY	1390	GTGCTTTTGAGGCTCATTTCTGGACGTCGACCCGTTGTATGCCAACAACTCTATGTAGAT	1449
Db	391	CTTCTTTATGAGATTATTTTCAGGAGAGTCTCTGTGCAATTATACCGGGCTCCGGGAGAA	332
QY	1450	GACAGCTTAGTTGACTCGGACACGACCATTTGCTTTAACCGAGCATCTGACGAAGGAGACTTT	1509
Db	331	GTGAATCTTGTGAATGGTTTGAAGAGATTGTGACGAATCGAGATCGGAAGGAGTTTG	272
QY	1510	GA 1511	
Db	271	GA 270	

RESULT 15

```

US-35-938-842A-1079
; Sequence 1079, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1079
; LENGTH: 2124
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1079

```

Query Match	9.9%	Score 192.4	DB 9	Length 2124
Best local similarity	55.0%	Pred. No. 3.4e-45		
Matches 476	Conservative 0	Mismatches 366	Indels 24	Gaps 4
QY	786	TTTCACATAGGAGGAGCTAGCTAGAGGCCACCAATGGTTTCCGAGGCCGAACCTGTGTAGG	845	
Db				
Db	1047	TCGTCTATGTAAGAAGCTCAAGAGGCACTAGCAAAATTTTGAATCTGTAGCAATCTAGG		
QY	846	ACAAGCGGGTTCGGTTACGTGCACAAAGSISGTTGCCCTAGTGGGAAAGAAATTGCTGT	905	
Db				
Db	1107	AGNAGTGGGTTTGGCAAGGTTTACAGAGCACTTTAGCGGATGGTACTCTCTAGCGAT	1166	
QY	906	GAAGCAGTGTAAAGTTGGGAGTGGTCAGGAGAGAGGGAGTTTCAGGCAGAGGTTTCAGAT	965	
Db				
Db	1167	TAAAGAGCTCACAAAGTGTGGGCCACCAAGGTGATAAGAAATTCAGGTCGAGATTGATAT	1226	
QY	966	CATCAGCAGAGTTCACCAAGGCATCTGGTGTCCTCTTTGGTTA-----TTTCATCGC	1019	
Db				
Db	1227	SGTTAGCCGCTCTCATCATCGTAATCTGTGAACCTTGTGGGTACTATAGTAGTCGAGA	1286	
QY	1020	CGGTGCCAAAGATTGCTTGTCTATCAGTTTGTCTCTACACAACTCTCGAGCTTCACCT	1079	
Db				
Db	1287	TTCTTCTCAGCACTACTTTGTATAGAGCTTGTCCAAATGGCAGCTCGAGCTTGGCT	1346	

QY	1080	CCATGG-----CGAGGACGGCCTCAATGGAAATGGAGCACCCAGATTTGAAGATTGCTCT	1133
Db	1347	CCATGGCCCTCTCGGTTGAACCTGCTCTTGATGGGACACCAGAATGAAGATTGCAC	1406
QY	1134	TGGATCGCTAAAGGACHTTCTTATCTTCATGAAGATTGCAATCCCTAAATCATTTCAACG	1193
Db	1407	TGATGCTGCAAGAGGACCTTCATACCTTCAIGAAGACTCGCAACCCCTCCGTTATACACAG	1466
QY	1194	TGATATCAAGGCTTCAACACATATGTAGATTTCAAGTTTGAAGCTAAGGTTCCTGATTT	1253
Db	1467	AGATTTTAAAGCCCTTAATATACCTCTTGAACACACTTCAAGCCCAAGTTGCAGATT	1526
QY	1254	TGCTCTGTCTAAGATTGCTTCTGATACAA---ACACGCAATGTATCAACACGTGTGATGG	1310
Db	1527	TGGCTTAGCCAAACAAGCTCCTCAAGCAGGSGTAATCACTTATCTACTGCTTTATGGG	1586
QY	1311	AACCTTTGGSTACTTTGGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAGNAGTCTGA	1370
Db	1587	CACATTTGGATATGTTGGCGCTGAATATGCAATGACGGGACACCTACTGCTCAAGAGTGA	1646
QY	1371	CGTTTCTCATTTGGCGGTGTGCTTTTGGAGCTCAITTACTGGACGTCGACCGGTGATGC	1430
Db	1647	TGCTTATAGTACGGTGTGGTCCCTCTCGAATTCGTAATGTTGGTAGAAMACCTTGAGNAT	1706
QY	1431	CAACAATGCTATGTAGATSCACAGTTAGTTGACTGGGCAACGACCAATTCCTTAACCGAGC	1490
Db	1707	GTCAACAACCTTCAGGCCAAGAAAUCTGCTCACTTTGGACAAGGCCAGTCTT-----	1757
QY	1491	AUTCAGCAGGAGACCTTTGAGGCTTTAGCTGATGCAAGATGAATATGGGNTGACAG	1550
Db	1758	RAGAGCAAGACCGGTTTAGAGACCTAGCTGCATTCACAGCTTGAGGAAAAATCCCGAA	1817
QY	1551	AGAGGAGATGGCTCGCATGGTTGCTGTGCTCGGCTTGTTGGCCATTCACCTCGCGCG	1610
Db	1818	AGAAGATTTCAAGAAGTATGCACAATCGCTGCAGCTTGTTGCACCTGAGCTAGCCA	1877
QY	1611	CAGACCTCCGATGAGCCGATTTGTC	1636
Db	1878	GAGACCAACGATGGCGGAAGTGGTTC	1903

Search completed: July 7, 2003, 08:04:54
Job time : 341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 07:29:23 ; Search time 2356 Seconds

(without alignments)
11023.819 Million cell updates/sec

Title: US-10-086-464-1

Perfect score: 1944

Sequence: 1 atgtctctggcgccgtctcc.....atagtgacctctctttaa 1944

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 3097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estl:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pla:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rtd:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442.2	22.7	759	13	BM408099 EST582426
2	429.6	22.1	757	12	BG596561 EST495239
3	413.2	21.3	692	12	BG441204 GA_Fa001
4	411.6	21.2	584	10	AV543493 AV543493
5	400	20.6	666	13	BM358715 GA_Fa001
6	399	20.5	573	10	AV551753 AV551753

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	376	19.3	1053	11	AY108241	Zea mays
8	366.2	18.8	1016	11	AY108243	Zea mays
9	358.4	18.4	645	14	BQ506869	EST614284
10	356.2	18.3	652	14	BQ849683	QGB10120
11	349.2	18.0	593	12	BF176907	EM1_4_B10
12	348.6	17.9	546	14	BQ240617	TaF05015B
13	348.2	17.9	537	14	BQ134241	BQ134241
14	345.8	17.8	679	14	BQ404121	GA_Ed006
15	337.2	17.3	770	10	BQ999193	QGG21G04
16	336.2	17.3	631	10	BE330924	SO9B061.Y
17	332	17.1	933	10	AW982539	HVSM9000
18	329.6	17.0	718	13	BM407268	EST581595
19	322.8	16.6	673	14	BQ410602	BM407268
20	320.4	16.5	608	14	BM814985	EST593079
21	316.8	16.3	561	10	AV539333	AV539333
22	315.2	16.2	643	12	BF203789	WHE1784.F
23	315	16.2	548	13	BM527398	sal61d03
24	308	15.8	593	14	BQ850206	QGB11P07
25	307.8	15.8	578	10	AV917322	AV917322
26	303.2	15.6	625	14	BM878778	PI6-C03.S
27	300.2	15.4	548	9	AI728857	BNLGH1118
28	299.4	15.4	484	12	BG817458	EM1_76_C0
29	299.2	15.4	573	14	BM815391	EST593485
30	298.4	15.3	497	12	BF201900	WHE1759-1
31	296	15.2	546	12	BG511034	sac65a03
32	295.4	15.2	529	10	BE637850	WHE1755-1
33	295	15.2	525	12	BG448469	BG448469
34	284.6	14.6	735	14	BQ788651	WHE4152.C
35	283.4	14.6	453	17	BH442220	BOGLN68TR
36	280.4	14.4	502	12	BG369831	HVSM91002
37	280.2	14.4	672	14	BQ005676	QGG8M18.Y
38	273.8	14.1	459	10	AV538962	AV538962
39	266.2	13.7	720	13	BC283891	BC283891
40	265.6	13.7	544	10	BE492937	WHE0563.B
41	258.6	13.3	489	10	AW982145	SS00075.S
42	256	13.2	628	10	AV826238	AV826238
43	255.6	13.1	618	10	AV826648	AV826648
44	254.4	13.1	610	13	BQ305865	BQ305865
45	253.6	13.0	436	10	AV793009	AV793009

ALIGNMENTS

RESULT 1
BM408099
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BM408099
EST582426 potato roots Solanum tuberosum cDNA clone cPRO33021 5'
end, mRNA sequence.
BM408099
BM408099.1 GI:18259729
EST.
potato.
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 759)
van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,
Tattersback,T., Chieming,A., Bougri,O., Buell,C.R., Rooning,C.,
Tanksley,S. and Baker,B.
Generation of ESTs from potato roots
Unpublished (2001)
Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: T3.
Location/Qualifiers
1..759
/organism="Solanum tuberosum"
/cultivar="Kennebec"

```

/db_xref="taxon:4113"
/clone="cPRO33021"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."
BASE COUNT      200 a 150 c 183 g 226 t
ORIGIN
Query Match      22.7%; Score 442.2; DB 13; Length 759;
Best Local Similarity 73.9%; Pred. No. 1.6e-95;
Matches 561; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
QY 933 GGGAGAGAGGAGTTTCAGGACAGAGTTGAGATCATCAGCAGAGCTTCACACAGGCATCT 992
Db 1 GGGGGAACGTGAATTCAGGCGGAAGTTGAGATTATTAGCGGATCATCATCAGCACT 60
QY 993 GGTGTCCTCTGTTGTTATTCGATCGCCGCTGCCAAAGATGCTGCTCATGAGTTGT 1052
Db 61 TGTGTCCTCTGTTGTTATTCGATCGCCGCTGCCAAAGATGCTGCTCATGAGTTGT 120
QY 1053 TCCCTAACCAACTCTCGAGCTTCACCTCCATGCGGAGGAGCGCTCATCATGAGTTGT 1112
Db 121 ICCAACAACTATTGGAATTCATTACAGCGAAGGAGGCGCTCTTGGATTGGCC 180
QY 1113 CACAGATGAAGATTCCTCTGGATCTGCTAAGGACTTCTTAICTTCATGAAGATTG 1172
Db 181 AATACGCTAAGATTGCTCTAGGTCAGCTAAGGAGCTGCTAICTGATGAGACTG 240
QY 1173 CATCTCTAATCATTCACGCTATATCAAGGCTTCAAGCTTCAAGATTGATGATTCAGTT 1232
Db 241 CCACCGAATCATTCACGCTATATCAAGGAGCTTCAAGCTTCAAGCTTCAAGTT 300
QY 1233 TGAAGCTAAGTTGCTGATTTGCTCTGCTAAGATTGCTCTGCTAACAACAGCATGT 1292
Db 301 TGAGGCTAAGTTGCTGATTTGCTCTGCTAAGATTGCTCTGCTAACAACAGCATGT 360
QY 1293 ATCAACAGCTGATGGGAACTTTGGTACTTGGCTCCGGAATAGCTGCAACCGGAAA 1352
Db 361 CTCCACAGAGTGATGGGAACTTTGGTACTTGGCTCCGGAATAGCTGCAACCGGAAA 420
QY 1353 GCTCAGGAGAGCTCTGAGCTTTCTCATTTGCGCTTGTGCTTTGGAGCTCATTTACTGG 1412
Db 421 GCTTACAGAGAGCTGAGCTTTCTCATTTGCGCTTGTGCTTTGGAGCTCATTTACTGG 480
QY 1413 AGCTGACCCGTTGATGCGCAACATGCTATGATGATGACAGCTTAGTTGACTGGCAGC 1472
Db 481 AGCTGCGCTGTTGACTCTACTCAATCATACATCGAGATAGTTGGTGGACTGGCAGC 540
QY 1473 ACATGCTTACCGAGCATCTGAGCAGAGAGACTTTGAGGGTTTACGCTGATGCAAGAT 1532
Db 541 TCCATTACTCACAGAGCTTTAGAGATGAAAAGTTGTAACCTTTGTTGATGCTCGGCT 600
QY 1533 GAATATGGTATGACAGAGAGAGATGCTCGCATGTTGCTTGTGCTGCGCTTGTGT 1592
Db 601 AGAATGATTTAACCATATGAGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1593 TCGCCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATTGCTGCTGCTGCTGCTGCTGCT 1652
Db 661 GCGTCACTCAGCAAGCGTAGCCAGCATGATGACAGGTTCTCCGAGGCTTGAAGGAGA 720
QY 1653 TGTATCAGTGTGATGCTTAAACGAAGGAGATGAGACGAG 1691
Db 721 TGTCTCATATACAGACCTTTACGAGGGATTAAACCTGG 759

```

RESULT 2

```

BG596561
LOCUS      BG596561          757 bp      mRNA      linear      EST 12-APR-2001
DEFINITION BG596561 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence,
            mRNA sequence.
VERSION    BG596561
KEYWORDS   EST.
SOURCE     potato.
ORGANISM   Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE  1 (bases 1 to 757)
            van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
            Bongri,O., Buell,C.R., Rouning,C., Tanksley,S. and Baker,B.
            Generations of ESTs from sprouting potato eyes
            Unpublished (2000)
            Contact: Cathy Renning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cda@resgen.com
            Seq primer: M13P-R.
FEATURES   Location/Qualifiers
            1..757
               /organism="Solanum tuberosum"
               /cultivar="Keonebec"
               /db_xref="taxon:4113"
               /clone="cSTS15A23"
               /clone_lib="cSTS"
               /tissue_type="sprouting eyes from tubers"
               /dev_stage="12-14 weeks post harvest"
               /lab_host="SOLR"
               /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
               XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
               taken from tubers. The tubers were incubated at 26C in the
               dark for 2-3 weeks prior to sprouting. The eyes were
               frozen in liquid nitrogen immediately upon removal from
               tubers."
BASE COUNT  201 a 149 c 184 g 223 t
ORIGIN

```

source

```

Query Match      22.1%; Score 429.6; DB 12; Length 757;
Best Local Similarity 73.4%; Pred. No. 1.7e-92;
Matches 549; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
QY 901 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGACAGGTT 960
Db 1 GCAGTTAAACAGCTTAAAGGCTGGAAGTGGACAGGGAACGTGAATTCAGGCGGAAGTT 60
QY 961 GAGATCATCAGCAGAGTTCCACAGCATCTGCTGCTCTCTGTTGTTATTCATCGCTC 1020
Db 61 GAGATTATTAGCGAGTACATCAGACATCTTGTGCTCTCTGTTGTTATTCATCGCTC 120
QY 1021 GGTGCCAAAAGATTGCTTGTCTATGATTTGCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
Db 121 GGGGCTCAGAGACTGCTGTTTATGATTTGTTTCCAAACAATACTTTGGAATTCATTTA 180
QY 1081 CATGGCGAGGAGCGGCTTCAATGATGATGAGACACAGATGAGATGATGCTTGGATCT 1140
Db 181 CACGGAAGGAGGAGCGCTTCTTGGATTGGCCATACGGCTAAAGATTGCTTAGGGTCA 240
QY 1141 GCTAAAGGACTTCTTATCTTCAATGAAGATTGCAATCTTAAATCATTCACCTGATATC 1200
Db 241 GCTAAGGACTGCTATCTGATGAGATTTCAAGTTTGAAGTTGCTGATTTTGGTCTT 1260
QY 1201 AAGGCTTCAACATATTGATGATTTCAAGTTTGAAGTTGCTGATTTTGGTCTT 1360
Db 301 AAGGAGCTAAATATCTTATCTGACTTTAATTTTGAAGCTAAGTTGCTGATTTGACTT 360
QY 1261 GCTAAGATTGCTTCTGATACAAACAGCATGTTTCAACAGCTGATGAGGAACTTTGGG 1320
Db 361 GCGAAGTAACTTCTGATGTTAATCTATCTCTCCACAGAGTGATGAGGAACTTTGGG 420

```



```
FEATURES
  source
    Location/Qualifiers
      1..584
        /organism="Arabidopsis thaliana"
        /strain="Columbia"
        /db_xref="taxon:3702"
        /clone="R220107F"
        /clone_lib="Arabidopsis thaliana roots Columbia"
        /tissue_type="roots"
        /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
          XhoI"
      BASE COUNT      143 a 154 c 213 g 174 t
      ORIGIN
        Query Match      21.2%; Score 411.6; DB 10; Length 584;
        Best Local Similarity 90.9%; Pred. No. 3.3e-88;
        Matches 438; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
      QY 1463 ACTGGGACGACCATGCTTAACCGAGCATCTGACCAAGGAGACTTTGAGGTTTACGTCG 1522
      Db 1463 ACTGGGACGACCATGCTTAACCGAGCATCTGACCAAGGAGACTTTGAGGTTTACGTCG 1522
      QY 1523 ATGCAAGATGAATATGGGTATGACAGAGGAGATGGCTCGCATGGTTGCTTCTGCTG 1582
      Db 1523 ATGCAAGATGAATATGGGTATGACAGAGGAGATGGCTCGCATGGTTGCTTCTGCTG 1582
      QY 1583 CGGCTGTGTCGCCATTCAGCTCGCCGACACCTCGCATGAGCCAGATTTGCGTGGCT 1642
      Db 1583 CGGCTGTGTCGCCATTCAGCTCGCCGACACCTCGCATGAGCCAGATTTGCGTGGCT 1642
      QY 1643 TAGAAGGAATATGATCACTGTCAGATCTTAACGAGGATGAGCAGGTCACCAAGCAATG 1702
      Db 1643 TAGAAGGAATATGATCACTGTCAGATCTTAACGAGGATGAGCAGGTCACCAAGCAATG 1702
      QY 1703 TATACAGCTCATACGAGGAAGACACCGATATGACTCGAGCCAGTACCAATGAAGACATGA 1762
      Db 1703 TATACAGCTCATACGAGGAAGACACCGATATGACTCGAGCCAGTACCAATGAAGACATGA 1762
      QY 1763 AGAAGTTAGAAAATGSCACTTGGAACTCAAGAGTACAAAGCCACGGGTGAGTACAGTA 1822
      Db 1763 AGAAGTTAGAAAATGSCACTTGGAACTCAAGAGTACAAAGCCACGGGTGAGTACAGTA 1822
      QY 1823 ATCCGACCAAGTACATGACTGTACCGCTCTGTTCAAGCAGCGAGGCGCAAAACACAC 1882
      Db 1823 ATCCGACCAAGTACATGACTGTACCGCTCTGTTCAAGCAGCGAGGCGCAAAACACAC 1882
      QY 1883 GCGAAATGGAGATGGGGAAGATTAAAGAAACCGGTCAAGGTTATAGTGGACCTTCTCTTT 1942
      Db 1883 GCGAAATGGAGATGGGGAAGATTAAAGAAACCGGTCAAGGTTATAGTGGACCTTCTCTTT 1942
      QY 1943 AA 1944
      Db 1943 AA 103

RESULT 5
BM358715
LOCUS
DEFINITION
  BM358715 666 bp mRNA linear EST 09-JAN-2002
  Gossypium arboreum 7-10 dpa fiber library Gossypium
  arboreum cDNA clone GA_Ea0012D16r, mRNA sequence.
ACCESSION
  BM358715
VERSION
  BM358715.1 GI:18099461
KEYWORDS
  EST.
SOURCE
  Gossypium arboreum.
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE
  1 (bases 1 to 666)
  Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
  D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  An integrated analysis of the genetics, development, and evolution
  of the cotton fiber
  Unpublished (2003)
```

COMMENT

Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases = 646
Seq primer: TAATAGACTCACTATAGG
High quality sequence stop: 666.
Location/Qualifiers

FEATURES

source

1..666
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0012D16r"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 178 a 134 c 167 g 187 t

BASE COUNT

ORIGIN

Query Match 20.6%; Score 400; DB 13; Length 666;
Best Local Similarity 75.6%; Pred. No. 2.1e-85;
Matches 496; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
QY 739 CTCTCCACCGCTCTCCAGGGCTTGTGTTAGGCTTCTCCAAAAGCACITTCACATACGAG 798
Db 739 CTCTCCACCGCTCTCCAGGGCTTGTGTTAGGCTTCTCCAAAAGCACITTCACATACGAG 798
QY 799 GAGTAGCTAGAGCACCACCAATGTTTCTCGAGGCGAAGTGTGTAGCAGAGCCGGTTC 858
Db 799 GAGTAGCTAGAGCACCACCAATGTTTCTCGAGGCGAAGTGTGTAGCAGAGCCGGTTC 858
QY 71 GAATTAGCAGAGCAACGATGCTTCTCGAAGTTAACCTCTTGGCAAGGTGGTTTT 130
Db 71 GAATTAGCAGAGCAACGATGCTTCTCGAAGTTAACCTCTTGGCAAGGTGGTTTT 130
QY 859 GGTTCAGTGCACAAAGGTGTGTTGCTAGTGGAAAGAGTGTCTGTGAGCAGTTTGAA 918
Db 859 GGTTCAGTGCACAAAGGTGTGTTGCTAGTGGAAAGAGTGTCTGTGAGCAGTTTGAA 918
QY 131 GGGTAGCTACACAAAGGAGTTCCTCCTAATGGAGAGTAGCAGZAAAGCACTCAAG 190
Db 131 GGGTAGCTACACAAAGGAGTTCCTCCTAATGGAGAGTAGCAGZAAAGCACTCAAG 190
QY 919 GTTGGAGTGTTCAGGAGAGAGGGAGTTTTCAGGAGAGTGTGAGATCATCAGCAGATT 978
Db 919 GTTGGAGTGTTCAGGAGAGAGGGAGTTTTCAGGAGAGTGTGAGATCATCAGCAGATT 978
QY 191 GCTGGAAGTGGGCAAGGCGAGAGAAATTCAGGCTGAAGTTGAGATCATTAAGCCGCTC 250
Db 191 GCTGGAAGTGGGCAAGGCGAGAGAAATTCAGGCTGAAGTTGAGATCATTAAGCCGCTC 250
QY 979 CACCACAGCATCTGGTGTCTCTTGTGTTATTCATCGCCGCTGCCAAGAGATGCTT 1038
Db 979 CACCACAGCATCTGGTGTCTCTTGTGTTATTCATCGCCGCTGCCAAGAGATGCTT 1038
QY 251 CATCACAAACCTCTGCTCTCAITGGTGGATACTGTATTTCTGGGACAAATAGATGCTT 310
Db 251 CATCACAAACCTCTGCTCTCAITGGTGGATACTGTATTTCTGGGACAAATAGATGCTT 310
QY 1039 GTCTATGAGTTTGTCTTAACCAACATCTCGACCTTCACCTCCATGGCGAGGAGCGCT 1098
Db 1039 GTCTATGAGTTTGTCTTAACCAACATCTCGACCTTCACCTCCATGGCGAGGAGCGCT 1098
QY 311 GTTATGAGTTTGTCTTAACCAACATCTCGACCTTCACCTCCATGGGAGGCGGACTG 370
Db 311 GTTATGAGTTTGTCTTAACCAACATCTCGACCTTCACCTCCATGGGAGGCGGACTG 370
QY 1099 ACAATGGATGGAGCAGCAGATGAGATGCTCTTGGATCTGCTAAGAGACTTCTTAT 1158
Db 1099 ACAATGGATGGAGCAGCAGATGAGATGCTCTTGGATCTGCTAAGAGACTTCTTAT 1158
QY 371 ACCATGGATGGGCGCAGAGGATGAAATTTGCTTAGGATCTSCAAAGAGACTTGGCATAT 430
Db 371 ACCATGGATGGGCGCAGAGGATGAAATTTGCTTAGGATCTSCAAAGAGACTTGGCATAT 430
QY 1159 CTTCATGAAGATTGCAATCTAAATCATTCACCGTGATATCAGGCTTCAACATATG 1218
Db 1159 CTTCATGAAGATTGCAATCTAAATCATTCACCGTGATATCAGGCTTCAACATATG 1218
QY 431 CTTCATGAAGATTGTCATCCTAAGATCATTCACCGTGATATTAAGCCCGCTATATCTG 490
Db 431 CTTCATGAAGATTGTCATCCTAAGATCATTCACCGTGATATTAAGCCCGCTATATCTG 490
QY 1219 ATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTTGGTCTTGTCTAAGATGCTTCTGT 1278
Db 1219 ATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTTGGTCTTGTCTAAGATGCTTCTGT 1278
QY 491 TTGGATTTCAAGTTTGAAGCAAAAGTGTCTGATTTTGGACTAGCGAAAATGCTTCCGAT 550
Db 491 TTGGATTTCAAGTTTGAAGCAAAAGTGTCTGATTTTGGACTAGCGAAAATGCTTCCGAT 550
QY 1279 ACAACACGATGATACACAGCTGTGATGGGAACCTTTGGTACTTGGCTCGGAAATAC 1338
Db 1279 ACAACACGATGATACACAGCTGTGATGGGAACCTTTGGTACTTGGCTCGGAAATAC 1338
QY 551 GTCACACGCGAGCTCTCCACCGAGGTGATGGGTACTTTTGGGTATTTAGCCCTCAGTAT 610
Db 551 GTCACACGCGAGCTCTCCACCGAGGTGATGGGTACTTTTGGGTATTTAGCCCTCAGTAT 610
QY 1339 GCTGCAAGCGGAAAGCTCACGAGAGAGTCTGACGTTTCTCATTTTGGGCTTGTGCT 1394
Db 1339 GCTGCAAGCGGAAAGCTCACGAGAGAGTCTGACGTTTCTCATTTTGGGCTTGTGCT 1394
QY 611 GCTTCAAGTGGAAAGCTCAGTATATATCAGATGTTTCTCTCTTCTGGGGTCAATGCT 666
Db 611 GCTTCAAGTGGAAAGCTCAGTATATATCAGATGTTTCTCTCTTCTGGGGTCAATGCT 666

```

RESULT 6
AV551753      573 bp      mRNA      linear      EST 06-SEP-2000
LOCUS        Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION   cDNA clone RZ13C07R 5', mRNA sequence.
ACCESSION    AV551753
VERSION      AV551753.1 GI:8723166
KEYWORDS     EST.
SOURCE       thale cress.
ORGANISM     Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE    1 (bases 1 to 573)
AUTHORS      Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE        A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL      DNA Res. 7, 175-180 (2000)
MEDLINE      20363093
COMMENT      Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     Location/Qualifiers
             1..573
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZ13C07R"
             /clone_lib="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"
BASE COUNT   147 a 132 c 139 g 155 t
ORIGIN
Query Match      20.5%; Score 399; DB 10; Length 573;
Best local Similarity 83.6%; Pred. No. 3.5e-85;
Matches 479; Conservative 0; Mismatches 85; Indels 9; Gaps 2;

QY      419  TGGAAATCGCATCGGAGGAGTCGCTCTGCTGTGATAGTACTGTGATTTGCTCTCT 478
Db      1   TAGGAATGCCATGAGGAGTGCCTATCTCTGTATACAGTCTGATTTGCTCTCT 60

QY      479  GTAAGAAGAAACACCGGAGA-----GAGGAAGAAGATGCTTACTATGTTCTCCGCCAC 532
Db      61  GTAAGAAGAAACGAAGAAGACAGCAGCATGAAGCTGCTTACTATGTTCTCTCTCT 120

QY      533  CTCTCTCTGTCCTCAAGCGCGAGGACCTTACGTTGGACAGCAGCACAATGCGGCAAC 592
Db      121  CTCATCTGGTCCCAAGCTGAGGAGCCTTACGGTGTCAACACACAGTAGTATGGCAACA 180

QY      593  AAAACGCAACACCAACCGCTCAGATCATGTGTCAGTCCTACTACCAACCAACCAACGCTC 652
Db      181  AAAACGCGTCACGGCGCTCAGATAATCATGTAGTACGTCATTTGCCACCACCTTAAGCTC 240

QY      653  CATCTCCACCAACGGACCTCTCT---CCACCTCCACCAACCGCTTTTCATGAGCAGCAGCG 709
Db      241  CATCTCCACCAACAAACCTCTCTCCGCCACCTCCACCAACCAACCAACGATTCAGTACG 300

QY      710  GCGGCTCCGACTACTCGGACCGTCCAGTCTCTCTCCACGCTCTCCAGGCGTTGCTTAG 769
Db      301  GTGGTCTGACTATTGGATCTTCGGTCTCTCTCCACATCTCCAGGCGTTGCTTAG 360

QY      770  GCTTCTCCAAAACACTTTCATACAGAGSAGCTAGCTAGAGCACCACCAATGGTTTCTCCG 829
Db      361  GCCTTTCTAAACCACTTTCATATGAGAGTGTGAGAGCTACTAATGGCTTCTCTG 420

QY      830  AGCGCAACTGTTAGGACAAGCGGGTTCGGTTACGTGCACAAAGTGTTCGCCTAGTG 889

```

Db 421 AGGCTAATTGTTTAGGACAGAGGAGGTTTGGTTAATGTCATAAAGGATPATATGCCCTAGTG 480
 Qy 890 GGAAGAAGATTGCTGTGAAGCAGCTTGAAGCTTGGAGTGGTCAAGGGAGAGAGGAATTTTC 949
 Db 481 GGAAGAAGATTGCTGTGAAGCAGCTTGAAGCTGGTAGTGGTCAAGGGAGAGAGACTTTC 540
 Qy 950 AGGCAGAGGTTGAGATCATCATCAGCAGATTCCAC 982
 Db 541 AGGCTGAGGTTGAGATCATCATCAGCAGATTCCATC 573

 RESULT 7
 AY108241
 LOCUS AY108241 1053 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PCO134818 mRNA sequence.
 ACCESSION AY108241
 VERSION AY108241.1 GI:21211319
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays.

 REFERENCE
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
 Arthur, L.W., Hanefey, M., Morgante, M. and Tincey, S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 AUTHORS 2 (bases 1 to 1053)
 COE, E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 FEATURES
 source
 1..1053
 /organism="Zea mays"
 /db_xref="MaizeDB:637889"
 /db_xref="taxon:4577"
 /clone="PCO134818"
 /clone_lib="Maize Mapping Project/DuPont Cornsensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"
 BASE COUNT 283 a 220 c 262 g 288 t
 ORIGIN

 Query Match 19.3%; Score 376; DB 11; Length 1053;
 Best Local Similarity 68.5%; Pred. No. 1.5e-79;
 Matches 566; Conservative 0; Mismatches 250; Indels 10; Gaps 3;

 QY 1104 GGAATGAGCAGCAGATTGAGATGCTCTTTGGAT-CTGCTAAGGACTTCTTATCTTC 1162
 Db 11 GGAGTGCCCTGCTAGATTAAGATCATGTTTGGTGCTGCCAAGGTTTAGCTTATCTTC 70

 QY 1163 ATGAAGATTGCAATCTCAATCATTCACCGTGATCATCAAGGCTTCAAAACATATTGATAG 1222
 Db 71 ATGAAGACTGCCATCCAAAGATCAICCATGTCACATAAAGGCATCTAAACATTCCTTCTTG 130

 QY 1223 ATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTTGCTTAAGATTGCTTCGTGATACAA 1282
 Db 131 ACTTCCAAATTTGAAGCTAAGTTGCTGATTTGGACTTGCAAAGTTTCACTACTGATTAACA 190

 QY 1283 ACAGCATGATCAACACGTTGATGGGACCTTTGGTACTTGGCTCCGGAATACGCTG 1342
 Db 191 ACACCCATGTTTCGACAAGAGTAATGGGACCTTTGGGTAATTTGGCACCTTGAGTAGCAT 250

 QY 1343 CAAGCGAAAGCTCACGGGAAGTCTGACGTTTTCTCATTTGGCGTTTGCTTTTGGAG 1402

```

Db      251  CTCTGGCAAGCTACAGAAATCCGATGTAATTTCTCTCGAGTCATGCTCTTGAGC 310
QY      1403  TCATTACTGGAGCTGCGACCGGTGATGCCAACAATGTCTATAGATGACACTTAGTTG 1462
Db      311  TTATTACTGGCGGCGACCGAGTTGACCAACCAACAATATATGATGACACTTGGTTG 370
QY      1463  ACTGGGCAAGCACTGCTTAACCGAGCATCTGACACAGGAGACTTTGAGGTITAGTTG 1522
Db      371  ACTGGGCAAGCCATTACTGATGCGAGCAGCTTGAGGATGGTGAATATGATGCTTATGG 430
QY      1523  ATGCAAGAAATGAATAATGGGTATGACAGAGAGAGATGGCTCGCATGTTGCTGCTG 1582
Db      431  ATCCCTCGSCTGGGAAGGACCTCAATCTAATGAGATGCGAAGATGATGAGCTGTGAG 490
QY      1583  CGGCTTGTGTCGCACTTACGCTCGCCGAGACCTCGCATGACGAGAGATGTTGCTGGT 1642
Db      491  CTGCTGTGTAGCCCATCTCGACGCTCGCGCCACGTATGAGTCAGTCTGCGGCTT 550
QY      1643  TAGAAGGAATGATCACTGTGAGTCTTAACGAGGAGATGAGCAGGCTCAAGCAATG 1702
Db      551  TGGAGGGAATGTTCTTTGGAGGACCTTAATGAGGTGTTGCGGCTCGCCATAGCCGT 610
QY      1703  TATACGCTCATAGCGGAAGCAACCGATTAATGACTCGAGCCAGTACAAATGAAGCATGA 1762
Db      611  TCTTTGGTCTATC---AGCAGCTCCGATTAAGTCTGCGCCAGTACACGAGGCATGA 667
QY      1763  AGRAGTTTAGAAATGGCACTTGAGCTCAAGAGTAGCAACGCCAGCGGTGAGTACAGTA 1822
Db      668  AGAAGTTCAAGAGATGGCACTT-----CAACAACAACATATACAGCAGCAATACAG 721
QY      1823  ATCCGACAGTACTATGAGTGTACCGCTCTGCTTCAAGCAGCGAGGCGCCAAACCCAC 1882
Db      722  CGCCAAACAGTGAATATGGACAGATACGCTTCTGATCAGCAGGAGGCGCCACGAGCG 781
QY      1883  CGGAATGAGAGTGGGAAGATTAAGAGAACCGGTGAGGTTATAG 1928
Db      782  AAGAGATGAGTGGTGGTGAATGAAGAAAGGTGGCTACAGTGTGG 827

```

```

RESULT 8
AV108243
LOCUS      Zea mays PC0134814 mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PC0134814 mRNA sequence.
ACCESSION AV108243
VERSION    AV108243.1 GI:21211321
KEYWORDS   HTC.
SOURCE     Zea mays.
ORGANISM   Zea mays

```

```

REFERENCE 1 (bases 1 to 1016)
AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
           Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
           Overgo Probes
JOURNAL    Unpublished (2002)
REFERENCE 2 (bases 1 to 1016)
AUTHORS    Coe,E.C.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (25-APR-2002) Maize Mapping Project, University of
           Missouri, Columbia, MO 65211, USA
FEATURES   Location/Qualifiers
            1..1016

```

```

            /organism="Zea mays"
            /db_xref="MaizeDB:637888"
            /db_xref="taxon:4577"
            /clone="PC0134814"
            /clone_lib="Maize Mapping Project/DuPont Consensus
            Library"
            /note="this sequence is part of a project of EST
            assemblies resulting from the application of public
            contigs to seed DuPont contigs; this resource was
            source

```

```

RESULT 9
BQ506869
LOCUS      BQ506869
DEFINITION EST614284 Generation of a set of potato cDNA clones for microarray
           analyses mixed potato tissues Solanum tuberosum cDNA clone STMG090
           5' end, mRNA sequence.
ACCESSION BQ506869
VERSION    BQ506869.2 GI:21922719
KEYWORDS   EST.
SOURCE     potato.

```

```

assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 261 a 251 c 262 g 262 t
ORIGIN

```

```

Query Match      18.8%; Score 366.2; DB 11; Length 1016;
Best Local Similarity 68.7%; Pred. No. 3.3e-77;
Matches 535; Conservative 0; Mismatches 238; Indels 6; Gaps 2;

QY      1115  CCAGATTGAGATTGCTCTTGATCTGTAAGGACTTTCTTATCTTCATGAGATTGCA 1174
Db      1  CTAGATTAAAGATCGCTCTGGGTGCTGCCAAGGGTTAGTTATCTTCATGAGACTGCC 60
QY      1175  ATCCCTAAATCATTCACCGCTGATATCAAGCTTCAACATATGATGATTTCAGTTTG 1234
Db      61  ATCCAAAGATCATCCACCGCGACATTAAGGCATCTAACATCTCTCTGTTGACTTCAAAATTG 120
QY      1235  AAGCTAAGGTTGCTGATTTTGTCTTGTCTTAAGATTGCTTCTGATACAAACACCATGTAT 1294
Db      121  AAGCTATGTTGCTGACTTTGGACTTGCAGATTCTACTACTGATACACACACCCATGTGT 180
QY      1295  CAACAGGTGTGATGGGAACCTTTGGTACTTGGCTCCGGAATACGCTCAAGCGGAAGGC 1354
Db      181  CAACAGAGTAAATGGGACCTTTGGTATTTGGCAACCGAGTATGAGCATCTGGCAAGC 240
QY      1355  TCACGGAGAAGTCTGACGTTTTTCATTTGGGCTTGTGCTTTTGGAGCTCATTAAGTGGAC 1414
Db      241  TCACAGAAATTCAGATGATTTTCTTCGGAGTCAATGCTTCTTGAGCTTATTAAGTGGG 300
QY      1415  GTGACCCGTTGATGCCAACAATGCTATGTAGATGACAGCTTAGTTGACTGGGCAACG 1474
Db      301  GCGCAGCAATTTGACAAACACCAACATATATGATGATGATGATGATGATGATGATGAT 360
QY      1475  CATTCCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGTTTACGTATGTCAGAAATGA 1534
Db      361  CATTAAGTATGAGAGCGCTCGAGGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY      1535  ATAATGGTATGACAGAGAGAGAGTGGCTGCGATGTTGCTGCTGCTGCTGCTGCTGCTG 1594
Db      421  GAAAGGACTTCAATCTAACAGAGATGCAAGATGATAGCTGTGCGGCTGCTGCTGCTGCT 480
QY      1595  GCCATTAGCTGCGCGAGACCTTCGCATGAGCCAGATTGTCGCTGCTGCTGCTGCTGCTG 1654
Db      481  GCCATTGCGCAGCTGCTGCGGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY      1655  TATCACTGTCAGATCTTAACGAAGGATGAGACCGAGTCAAGCAATGATATACACTCAT 1714
Db      541  TGTCTTTGGAGACCTTATGAGGTGTTCCAGCTTGGCCATAGCGGCTTCTTTGGGTGAT 600
QY      1715  ACGGAGGAAGCACCAGTATGACTCGAGCCAGTACATGAAGATGAAGATGAAGATGAAG 1774
Db      601  A---TAGCAGCTCTGATTACGATTCGCGGAGTACACAGGAGGACATGCAAGATTCAGGA 657
QY      1775  AAATGGCAGTGGTACTCAGAGTACACGCCCGGTGAGTACAGTATTCGACCAAGT 1834
Db      658  AGACGGGCTT---CAATAACAACACTTACACGAGCCATACATACAGCGCGCCACCAAGT 714
QY      1835  ACTATGACTGTACCGCTTGGTTCAAGCAGGAGGCGCAACCAACACAGCGGAATGAG 1893
Db      715  AATACGCCAGGTGCGCTGCTGGATCAAGCAGGAGGCGCGCCGACGACGAGGATGAG 773

```

ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 645)
 AUTHORS Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J., and
 Karamycheva, S.A.
 TITLE Generation of a set of potato cDNA clones for microarray analyses
 JOURNAL Unpublished (2002)
 COMMENT On Jun 10, 2002 this sequence version replaced gi:21365738.
 Other ESTs: EST614285
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato@tigr.org
 This clone is available through the Research Genetics, contact the
 Research Genetics for further information 1-800-711-6195 or
 cda@resgen.com
 Seq primer: T3.
 FEATURES
 source Location/Qualifiers
 1..645
 /organism="Solanum tuberosum"
 /cultivar="Kennebec or Binjite"
 /db_xref="taxon:4113"
 /clone="STM390"
 /clone_lib="Generation of a set of potato cDNA clones for
 microarray analyses mixed potato tissues"
 /tissue_type="mixed tissues"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
 XhoI; supplier: Combination of untreated and Phytophthora
 infestans-treated libraries of stolons, leaves, leaflets,
 axillary buds of stem explants, petioles, germinating eyes
 , tubers, or roots."
 BASE COUNT 175 a 124 c 154 g 192 t
 ORIGIN
 Query Match 18.4%; Score 358.4; DB 14; Length 645;
 Best Local Similarity 74.3%; Pred. No. 2.1e-75;
 Matches 452; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
 QY 901 GCTGTGACGAGCTGAAAGTTGGAGTGTGACGAGGAGAGAGGAGTTTCAGGACAGAGTT 960
 DB 1 GCAGTTAAACAGCTTAAGCTTGGAGTGGAGGAGGAGGAGGAGTTTCAGGCGGAGATT 60
 QY 961 GAGATCATCAGCAGAGTTTCACACAGGCACTGTGTCTCTGTGTGTATGTCATCC 1020
 DB 61 GAGATTATTAGCGGATACATCACAGCACTGTGTCTCTGTGTGTATGTCATCTACT 120
 QY 1021 GTTGCCAAAGATGCTGTCTATGAGTTTGTCTTACACACAACTCTGAGCTTCACCTC 1080
 DB 21 GGGCTCAGAGACTGCTGTATGAGTTTGTCTTCCAAACAACTCTTGGAAFTTCATTTA 180
 QY 1081 CATGCGGAGCGGCTACAAATGGAATGGAGCACCAGATTCAGAGTTGCTCTTGGATCT 1140
 DB 181 CACGGAAGGAGGAGGCTCTTGGATTGGCCATACGCTAAAGATTGCTCTAGGTCAC 240
 QY 1141 GCTAAGGAGCTTTCTTAFTCTCATGAGAAITGCAATCCATAATCATTACCGTGATATC 1200
 DB 241 GCTAAGGAGCTGGCATATCTGCTGATGAGACTGCCAACCGAAATCATTACCGTGATATC 300
 QY 1201 AAGCTTCAACATATTGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTTGGCTT 1260
 DB 301 AAGCAGCTAACTACTTATCGACTTTAATTTGAGGCTAAGTTGCTGATTTTGGACTT 360
 QY 1261 GCTAAGATTGCTTGTGATAAACAACGCGATGATCAACAGCTGATGGAGACCTTTGGG 1320
 DB 361 GCCAAGCTAACTCTGATGTTAATTAATCTCATGCTCCACAGAGTATGGAGACCTTTGGG 420
 QY 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTTCACGGAGAGTCTGACGTTTCTCA 1380
 DB 421 TATTTGGCTCCAGATATGCTTCTCTGGAAGCTTACAGACAAGTACAGACGATTTCTCC 480

QY 1381 TTTCGCGTGTTCCTTTGGAGCTCAATTACTGAGCTGACCGCTGATGACCAACATGTC 1440
 DB 481 TTGTGTAATGCTTCTTTGAGTTGATTAAGTGAAGCTGCGCTGTGACTCTACTCAATCA 540
 QY 1441 TATGTAGATGACAGCTTACTTACTGCTGGGACAGCAATCTCTTAACCGAGCATCTGAGCAA 1500
 DB 541 TATATCGAGATAGTTTGTGTGACTGGGACGCTCCATTAATCTACACAGGACTTAGAAGAT 600
 QY 1501 GGAGACTT 1508
 DB 601 GAAAGTTT 608
 RESULT 10
 BQ849683
 LOCUS
 DEFINITION BQ849683 652 bp mRNA linear EST 14-AUG-2002
 QGB10120.yq.abl OG.ABCDI lettuce salinas Lactuca sativa cDNA clone
 QGB10120, mRNA sequence.
 ACCESSION BQ849683
 VERSION BQ849683.1 GI:22235152
 SOURCE EST.
 ORGANISM Lactuca sativa.
 Lactuca sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
 Lactuca.
 REFERENCE 1 (bases 1 to 652)
 AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison,
 P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, L. and Bradford, K.
 TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://comgenomics.ucdavis.edu/
 JOURNAL Unpublished (2002)
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmudson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.org [michelmore@vegmil.ucdavis.edu]
 belongs to contig QG_CA.Contig2262, see http://cgdb.ucdavis.edu/
 for details.
 Plate: QGB10 row: I column: 20.
 Location/Qualifiers
 1..652
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QGB10120"
 /lab_host="E.coli"
 /note="Vector: pBRCDNASFIAB; The library was constructed
 from 10 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgdb.ucdavis.edu/
 TAG_LIB=QG.ABCDI lettuce salinas
 TAG_TISSUE=roots
 TAG_SEQ=GTTCACGGG"
 BASE COUNT 175 a 127 c 157 g 193 t
 ORIGIN
 Query Match 18.3%; Score 356.2; DB 14; Length 652;
 Best Local Similarity 71.8%; Pred. No. 7.1e-75;
 Matches 466; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

REFERENCE	1. (bases 1 to 646)
AUTHORS	Cloutier, S.
TITLE	Wheat functional genomics - Glenlea developing seeds cDNA libraries
JOURNAL	Unpublished (2002)
COMMENT	Contact: Dr. Sylvie Cloutier Cereal Research Centre, Agriculture and Agri-food Canada 195 Dafoe Rd. Winnipeg, MB, Canada R3T 2M9 Tel: (204) 983-2340 Fax: (204) 983-4604 Email: scloutier@agr.ca was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end). Average insert size is >2.0 kb Plate: 015 row: B column: 02 Seq primer: M13 Reverse.

```

FEATURES             source
seq primer: M13 Reverse-
Location/Qualifiers
1..646
    .organism="Triticum aestivum"
    /cultivar="Glenlea"
    /db_xref="taxon:4365"
    /clone="TaE05015B02R"
    /clone_lib="TaE05"
    /tissue_type="developing seeds"
    /dev_stage="5 days after anthesis"
    /lab_host="E. coli DH10B"
    /note="Vector: pSPORT-P (Invitrogen Technologies); Site_1:
NotI; Site_2: MluI; mRNA obtained from wheat seeds of
cultivar Glenlea 5 days post-anthesis"
163 a      126 c      162 g      189 t

```

Query Match	17.9%	Score 348.6	DB 14	Length 646
Best Local Similarity	71.8%	Pred. No. 4.8e-73		
Matches 456	Conservative 0	Mismatches 179	Indels 0	Gaps 0
QY	957	GGTGGAGATCATCAGCAGAGTTCACACACAGGCATCTGGTGCTCTCTCTGTTGGTTATGTCAT	1016	
Db	12	GGTGGAGATTCACGCCGGTTCATCACAAGCATCTGTGCTCTCTCTCTGTTGGTTATGTCAT	71	
QY	1017	CGCGGTGCCAAAGAATCTCTGTCTATGAGTTTGTTCTCTTAAACAANTCTCAGAGTTCA	1076	
Db	72	CTCTGGGGGGAAGAGGTTCGTTGTATGATGTTGTCCACCAATAACACATGGAATCCA	131	
QY	1077	COTCATGCGGAGGAGCGCTACATGCAATGGAGCACCAGATTCGAAGATTCCTCTGG	1136	
Db	132	CTTACATGGAAGAAGCCGTCACAGTGGAGTGGCCCATGAAGACTAAGGATGCCCTTGG	191	
QY	1137	ATCTGCTAAAGACATCTCTTCAATGATGAAGATTGCAATCTCTAAATCATTCACCGTGA	1196	
Db	192	TGCTGCTAAGGGTTTGGCATACATTCATGAAGATTGCCACCGAAGATCAIACATCGTGA	251	
QY	1157	TATCAAGGCTTCAAAACATATTCATGATTTTCAAGTTTGAAGCTTGAAGTTCTGTGAITTTGG	1256	
Db	252	TATAANGTCATCAACATCTCTTGTATTTTAATTCGAAGCTAAGGTTGCGGAHTTTGG	311	
QY	1257	TCTTCTAAGATTGCTTCGTATCAACAACCGCATGTATCAACACCGTGTGATGGGAACCTT	1316	
Db	312	TCGTGCAAGATTACCTCTGTATCAACAACACACATGTCTCAACAAGAGTAATGGGCACCTT	371	
QY	1317	TGGGTACTTGGCTCCGGGAATACGCTGCAAGCGGAAGCTCAGGAGAAGCTCTGACGTTT	1376	
Db	372	TGGGTATCTAGCACCCAGAAATGCGCTCTCTGGAAGCTAACTGAGAATCAGATGCTCT	431	
QY	1377	CTCATTTGGCGTTCGCTTTTGGAGCTCATTTACTGGAGCTGCACCGCTTGATGCCAACAA	1436	
Db	432	TTCCCTTTGGAGTGAATGCTTCTTTGAGCTGATTAACCTGGGGTCCGCGCTGTGATTCGACCCA	491	
QY	1437	TGCTATGTAGATGACAGCTTAGTTGATCTGGCAGGACCATTTGCTTAACCGAGCATCTGA	1496	
Db	492	AACATATAIGGATGACACCTTGGTTGATTTGGCAAGACCTTTTACTGTATCGGAGCACTTGA	551	
QY	1497	GCAAGGACATTTGAGGGTTTAGCTGTATGTCAAAGAAGAAATAGTGGGTATGACAGAGAGA	1556	

552 GGATGGTAACATGATGAGTTAGTGCGTCTGCGAAGAGGATTTCAATCAATGA 611
 QY 1557 GATGGCTCCGATGTTGCTGTGCTCGGCTGTG 1591
 DB 612 GATGCCAGATGATAGCATGTGCTGCAATGG 646

 RESULT 13
 BQ134241
 LOCUS
 DEFINITION 1091014D01.y2.1091 - Immature ear with common ESTs screened by
 Schmidt lab Zea mays cDNA, mRNA sequence.
 ACCESSION BQ134241
 VERSION BQ134241.1 GI:20208152
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 637)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 UNIVERSITY
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1091014 row: D column: 01.
 Location/Qualifiers
 1. .637
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="1091 - Immature ear with common ESTs screened
 by Schmidt lab"
 /tissue_type="Inflorescence meristem - floral organ
 primordia"
 /dev_stage="0.5 cm to 2 cm"
 /lab_host="Stratagene XL0LR"
 /note="Organ: Immature ear; Vector: pAD-GAL4; Site.1:
 EcoRI; Site.2: XhoI; RNA from library 606 was filtered for
 common ESTs found in 606."
 BASE COUNT 165 a 133 c 156 g 183 t
 ORIGIN
 Query Match 17.9%; Score 348.2; DB 14; Length 637;
 Best Local Similarity 73.2%; Pred No. 5.9e-73;
 Matches 446; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
 QY 712 GGCTCGGACTACTCGGACCGECCAGCTTCCTCCACCGTCTCCAGGGCTTGTGTAGGC 771
 DB 29 GGTTCCTAATTACTCTGGTGGCGAGATCCTACCTCCACCATCCCTGGTGCCTGTC 88
 QY 772 TTCTCCAAAGCAGCTTTCATACACGAGAGCTAGCTAGGCCACCACTGTTTCCTCGAG 831
 DB 89 TTCTCGAGAGCAGCAATTCACCTACGAGAGCTGTTGAGGGCTACTGATGGATTCGGAT 148
 QY 832 GCGAACTGTTTAGGACAAAGGGGGTTCGGTTACGTGCACAAAGGTGTCTGCCAGTGGG 891
 DB 149 GCTAATCTCCTGGACAGGTGGTTTCTGTTATGTTCCATATGATTCCTGCCAATGGC 208
 QY 892 AAAGAAGTGTGTGAAGCAGTTGAAGTTCGGGAGTGGTCCAGGAGAGGAGGATTTCCAG 951
 DB 209 AAAGAGATTGCTGTAAACAATTGAACCTGGGAGTGGCCAGGAGAGCGGTGATCCAG 268
 QY 952 CGCAGAGTTGAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGTCTCTTTGTTGTTAT 1011
 DB 269 GCTGAGTGTGAGATTATCAGCGAGTATCATCAAAACACCTTGTGCTTTGGTGGCTAT 328

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmunsdon Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
 belongs to contig QG_CA_Contig2262, see http://cgdb.ucdavis.edu/
 for details.

Plate: QG621 row: G column: 04.

FEATURES
source

Location/Qualifiers
 1..770
 /organism="Lactuca sativa"
 /cultivar="L. serriola"
 /db_xref="taxon:4236"
 /clone="QG621G04"
 /clone_lib="QG_EFGHJ lettuce serriola"
 /lab_host="E.coli"

/note="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_LIB-QG_EFGHJ lettuce serriola
 TAG_TISSUE=chemical induction
 TAG_SEQ=TAGCGCGG"

BASE COUNT 197 a 201 c 177 g 194 t 1 others
 ORIGIN

Query Match 17.3%; Score 337.2; DB 14; Length 770;
 Best Local Similarity 69.5%; Pred. No. 2.8e-70;
 Matches 485; Conservative 0; Mismatches 209; Indels 4; Gaps 2;

```

QY 625 ACCTACTACACACACACCTTAAGGCTCCATCTCCACACGCAACCTCCCTCCACCTCCA 684
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 ACTCTCCACACACACCTCGGAGACCTATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 685 CCACGGCTTTTCATGAGACAGACGCGGGTCCGACTACTCGGACCGTCCAGTTCCTCT 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 CCGGCCCTATTTAACAGCGGTGTTCGGTTCCAACTATTCCGGCGTTCAACCTCTT 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 745 CCACCGTCT---CCAGGCTGTGTAGGCTTCCCAAAKSCACTTTCACATCAGGAG 801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 CCACCGCTTCCCGACGAATGCTACTGGGTTTCTCAAGAGCACATTCACATCAGAGAA 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 802 CTAGCTAGACCCACCAATGGTTTCTCCAGCGAACTTGTITAGGACAGCGGGTTCGGT 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 TTACGAGAGCCACAGATGGTTTCCGAGCCAAATTTCTCGCACAGGTGGTTTGGT 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 862 TAGTGCACAAAGGTGTGTCCTAGTGGGAAGAAGTTCGTGTGAAGCAGTTGAAAGTT 921
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 314 TAGTCCACCGTGGTGTCTTCGGAACGGGAAGAGGTGGCGGTGAAGCACTAAAAACG 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 922 GGGAGTGGTCAGGAGAGAGGGAGTTTCAGGACAGGTGTGAGTCAATCAGCAGAGTTAC 981
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 GGCAGCGGGCAGGGGAGAGTGTGTTTCAGCGGAGGTGAGATCATATAGTCGAGTTAT 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 982 CACAGGCATGCTGTGCTCTGTGTGTATTCATCGCGTCCCAAAAGATTGCTGTC 1041
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 CATAAACATCTCGTTCTTTTAGTTGGATATTGCATATCAGAGCTCAAGATGGTGT 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1042 TATCAGTTTGTCTTACACATCTCGAGCTTCACCTCCATCGGAGGAGCGGCTTACA 1101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 TATGAGTTTGTTCCTCAATACACCAATGGAATCCACTTACATGAAAGATTCGTCGGTA 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1102 ATGGAATGGAGACACAGATTGAGATTGCTCTTGGATCTCTAAAGACTTTCTATCTT 1161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 ATGGAGTTTCCCAAGGCTGGCAATGTCTTTGGAGCTGCCAAGAGACTCGCGTATCTC 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: July 7, 2003, 10:26:51
 Job time : 2874 secs

```

QY 1162 CATGAAGATGGAATCCTAAATCATTCACCGTGATATCAAGGCTTCAACATATTGATA 1221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 CATGANGATGSCCATCCGAAGATCATTCATCGAGATATAAAAGCTGCTAATATACTCTT 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1222 GATTTCAGTTTGAAGCTAAGGTGCTGATTTTGCTTCTAGATTCCTCTCTGATGACA 1281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 GATTTCAGTTTGAAGCAAGGTAGCGGATTTGGGCTTGCAGAGATACTTCCGAT-GT 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1282 AACACGCATGATACACACACGCTGTGATGGGAACCTTTGG 1319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 733 GGCACATCATGCTCTCGACCCCGTGTGATGGGAAGTTTGG 770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 2, 2003, 15:16:17 ; Search time 25 Seconds
(without alignments)
1073.407 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSPGTEPPSPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	656.5	19.0	410	1	APKA_ARATH
2	654.5	19.0	412	1	APKB_ARATH
3	647.5	18.8	942	1	TMKL_ARATH
4	636.5	18.1	389	1	NAK_ARATH
5	605	17.5	901	1	CR14_MAIZE
6	579.5	16.8	390	1	CLVL_ARATH
7	572	16.5	849	1	SRK6_BRAOL
8	525	15.2	399	1	RLK5_ARATH
9	515.5	14.9	710	1	IRAL_MOUSE
10	512.5	14.8	817	1	KPRO_MAIZE
11	500	14.5	712	1	IRAL_HUMAN
12	434	12.6	555	1	GPI_CHLRE
13	374	10.8	674	1	TMKL_ARATH
14	367.5	10.6	501	1	KPEL_DROME
15	360	10.4	620	1	EXTN_TOBAC
16	355	10.3	296	1	PRP3_MOUSE
17	353	10.2	1490	1	CRK7_HUMAN
18	348	10.1	913	1	DDRL_HUMAN
19	346.5	10.0	267	1	EXTN_MAIZE
20	342	9.9	5179	1	MOC2_HUMAN
21	337	9.8	261	1	PRP2_MOUSE
22	332	9.6	426	1	EXLP_TOBAC
23	329.5	9.5	807	1	RMIL_COTJA
24	329.5	9.5	1584	1	KYKL_DICDI
25	328.5	9.5	743	1	EXT2_ARATH
26	326.5	9.5	911	1	DDRL_MOUSE
27	325	9.4	806	1	RMIL_CHICK
28	324.5	9.4	910	1	DDRL_RAT
29	323.5	9.4	283	1	EXTN_SORBI
30	320	9.3	1174	1	KPCL_COCHE
31	319	9.2	590	1	IRAZ_HUMAN
32	318.5	9.2	1386	1	ZAP3_MOUSE
33	316	9.2	503	1	WAIP_HUMAN

ALIGNMENTS

RESULT 1

APKA_ARATH
ID APKA_ARATH STANDARD; PRT; 410 AA.
AC 006548;
DI 01-NOV-1995 (Rel. 32, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein Kinase APK1A (EC 2.7.1.-).
GN APK1A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94081726; PubMed=1450380;
RA Hirayama T., Oka A.;
RT "Novel protein kinase of Arabidopsis thaliana (APK1) that phosphorylates tyrosine, serine and threonine.";
RL Plant Mol Biol. 20:653-662(1992)
CC -!- FUNCTION: POSSIBLE BI-FUNCTIONAL KINASE. IN VITRO, IT EXHIBITS SERINE/THREONINE ACTIVITY. IN VIVO, CAN PHOSPHORYLATE TYROSINE RESIDUES OF LIMITED SUBSTRATES.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; D12522; BAA02092.1; -;
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; Multigene family; Myristate.
CC LIPID 2 2 MYRISTATE (BY SIMILARITY).
CC DOMAIN 68 352 PROTEIN KINASE.
CC NP_BIND 74 82 ATP (BY SIMILARITY).
CC BINDING 106 106 ATP (BY SIMILARITY).
CC ACT_SITE 203 203 BY SIMILARITY.
CC SEQUENCE 410 AA; 45519 MW; 5B8B28D9E0065082 CRC64;

Query Match

Best Local Similarity 39.1%; Score 636.5; DB 1; Length 410;

Matches 160; Conservative 63; Mismatches 131; Indels 55; Gaps 12;

QY 238 GSDSYDRPVLPPP-----SPGLVLFESKSTIETEEELARATNGFSEANLGGCGFYV 289

Db 29 GSKASSVYRSPRTEGEILQSPNL-----KSFSAELKSAFNRPRDSVLGEGFGCV 82

QY 290 HGVLP-----SKGEVAVKLVGGGCGERFQAEVEIISRVHRHLVSLVGYCI 339

Db 63 FKGWIDEKSLTASRPTGLVIAVKKLNDQWGHQEWLAENVYLGQFSRHLVKLIGYCL 142

QY 340 AGAKRLVYEEFVNNLEHLHGG--RPTMENSTRKIALGSAKGLSYLHEDCNPKII 396

Db 143 EDEHLVYEEFVNNLEHLHGG--RPTMENSTRKIALGSAKGLSYLHEDCNPKII 396

QY 397 HRDIKASNLIDFKFEAKVADFLAKIAS--DNTVSTRVMGTGFLAPEYLAASGLTK 455

Db 201 YRDEKTSNILLDSYNAKLSDFGLAKDGPIGOKSHVSVMTGTHYAAPEYLAHGLHTK 260

QY 456 SDVTFSGVYLLITGRPRVDANNVYVDSLVDMARPLNRAEQDGFGLADAKNNGY 515

Db 261 SDVTFSGVYLLITGRPRVDANNVYVDSLVDMARPLNRAEQDGFGLADAKNNGY 515

QY 516 DREEMARVACAAACVRSARRPRMSQIVRALEGNVSLDLNEGRPCQSNVSYSGS 575

Db 318 SMEZACKVATLSRLCTTEIKLRPNMSEVSVSHLEHIOSLN-----AAIGN 363

QY 576 TDYDSQY--NEDMKFRKMGALGTQETNATGEYSNPTSDYGLYPSGSS 622

Db 364 MDKIDRRMRSSDSVSKVAGFAQTAVG-----STVVAYPRPAS 406

RESULT 2

APKB_ARATH STANDARD; PRT; 412 AA.

AC 01-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein kinase APK1B (EC 2.7.1.3.-)

GN APK1B OR AT2G28930 OR T914.1.

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_taxid=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X, Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,

RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,

RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,

RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,

RA Niernan W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

RA Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

RT thaliana."

RL Nature 402:761-768(1999).

RN [2]

RP SEQUENCE OF 143-346 FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=93081726; PubMed=1450380;

RA Hirayama T., Oka A.;

RT "Novel protein kinase of Arabidopsis thaliana (APK1) that

RT phosphorylates tyrosine, serine and threonine."

RL Plant Mol. Biol. 20:653-662(1992).

CC -!- FUNCTION: POSSIBLE BI-FUNCTIONAL KINASE. IN VITRO, IT EXHIBITS

CC SERINE/THREONINE ACTIVITY. IN VIVO, CAN PHOSPHORYLATE TYROSINE

CC RESIDUES OF LIMITED SUBSTRATES (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

OS Arabidopsis thaliana (Mouse-ear cross).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AC005315; AAC33221.1; --

DR EMBL; D10152; BAA20968.1; --

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR004040; STY_pkinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00089; pkinase; 1.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00221; STYKG; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;

KW ATP-binding; Multigene family; Myristate.

FT LIPID 2 2 MYRISTATE (BY SIMILARITY).

FT DOMAIN 69 356 PROTEIN KINASE.

FT NP_BIND 75 83 ATP (BY SIMILARITY).

FT BINDING 107 107 ATP (BY SIMILARITY).

FT ACT_SITE 204 204 BY SIMILARITY.

SQ SEQUENCE 412 AA; 45746 MW; EB1CA0B1A626A5DA CRC64;

Query Match 19.0%; Score 654.5; DR 1; Length 412;

Best Local Similarity 41.3%; Pred. No. 1.7e-18;

Matches 152; Conservative 62; Mismatches 113; Indels 41; Gaps 11;

QY 226 PPPPPFMSS-----SGGSDYSRPLPPP-----SPGLVLFESKSTIETEEELARATN 273

Db 14 PGASPKYMSSEANDSLGSKSSSVSINTNPTEGEILQSPNL-----KSFSAELKAAAF 67

QY 274 GFSEANLLGCGFGYVHKGVLP-----SGKEVAVKLVGGGCGERFQAEVEI 323

Db 68 NFRPDSVLGEGFGSGVFKGWIDEQTLTASKGIVGVIAVKKLNDQWGHQEWLAENVYL 127

QY 324 SRVHRHLVSLVGYCIAGAKRLVYEFVNNLEHLHGG--RPTMENSTRKIALGSA 381

Db 128 GQFSHPNLVKLIGYCLEDEHRLVYEFVNNLEHLHGG--RPTMENSTRKIALGSA 187

QY 382 KGLSYLHEDCNPKIIHRDIKASNLIDFKFEAKVADFLAKIA--SDNTNTHVSTRVMGTG 440

Db 188 KGLAFLH-NAETSVIYRDFKTSNILLDSYNAKLSDFGLAKDGPITGKSHVSTRMGTYG 246

QY 441 YLAPEYLAASGLTKESDVFSGVLELITGRPRVDANNVYVDSLVDMARPLNRAEQD 500

Db 247 YAAPEYLAHGLHTKSDVYSYGVLLVLEISGRVADKRRPPEGEQKLVEKARPLL--ANKR 304

QY 501 GDFEGLADAKNNGYDREEMARVACAAACVRSARRPRMSQIVRALEGNVSLDLNEGR 560

Db 305 KLER-VIDNRLQDQYSMEACKVATLALRCITFEIKLRPNMSEVSVSHLE---HIOTLNEA 360

QY 561 MRPGQSNV 568

Db 361 ---GGRNI 365

RESULT 3

TMKL_ARATH STANDARD; PRT; 942 AA.

ID TMKL_ARATH

AC P43298;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Putative receptor protein kinase TMK1 precursor (BC 2.7.1.1.-).

GN TMK1 OR AT1G36150 OR F15E12.4.

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NBL_TaxID=3702;

RN [1];

RN SEQUENCE FROM N.A.

RN STRAIN=cv. Columbia;

RX MEDLINE=93076110; PubMed=1332795;

RA Chang C., Schaller G.E., Patterson S.E., Kwok S.F.,

RA Meyerowitz E.M., Bleecker A.B.;

RT "The TMK1 gene from Arabidopsis codes for a protein with structural

RT and biochemical characteristics of a receptor protein kinase.,"

RL Plant Cell 4:1263-1271(1992).

RN [2];

RN SEQUENCE FROM N.A.

RN STRAIN=cv. Columbia;

RX MEDLINE=21016719; PubMed=11130712;

RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,

RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,

RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,

RA Dunn P., Etgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujil C.Y.,

RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,

RA Hurter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,

RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,

RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li X.-P.,

RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,

RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,

RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,

RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,

RA Sur H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,

RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;

RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis

RT thaliana.,"

RL Nature 408:816-820(2000).

CC -!- FUNCTION: PROBABLE RECEPTOR.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.

CC -!- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC -!- SIMILARITY: CONTAINS 9 LEUCINE-RICH REPEATS (LRR).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L00670; AAA32876.1; --

DR EMBL; AC026480; AAG51302.1; --

DR HSSP; P12931; IPMK.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003592; LRR_out.

DR InterPro; IPR003591; LRR_tyr.

DR InterPro; IPR004040; STY_pkinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00369; pkinase; 1.

DR Pfam; PF03560; LRR; 11.

DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00370; LRR; 7.

DR SMART; SM00369; LRR_TYR; 1.

DR SMART; SM00221; STYK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR Transferase; Serine/threonine-protein kinase; ATP-binding;

DR Transmembrane; Receptor; Glycoprotein; Signal; Repeat;

DR Leucine-rich repeat; Phosphorylation.

DR SIGNAL 1 23

DR POTENTIAL.

FT CHAIN	24	942	POTATIVE RECEPTOR PROTEIN KINASE TMK1.
FT DOMAIN	24	482	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	483	503	POTENTIAL.
FT DOMAIN	504	942	CYTOPLASMIC (POTENTIAL).
FT REPEAT	87	110	LRR 1.
FT REPEAT	111	133	LRR 2.
FT REPEAT	135	159	LRR 3.
FT REPEAT	185	209	LRR 4.
FT REPEAT	231	253	LRR 5.
FT REPEAT	254	278	LRR 6.
FT REPEAT	280	300	LRR 7.
FT REPEAT	385	408	LRR 8.
FT REPEAT	409	436	LRR 9.
FT DOMAIN	589	869	PROTEIN KINASE.
FT NP_BIND	594	602	ATP (BY SIMILARITY).
FT BINDING	616	616	ATP (BY SIMILARITY).
FT ACT_SITE	717	717	BY SIMILARITY.
FT CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	99	99	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	164	164	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	171	171	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	230	230	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	285	285	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	363	363	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	397	397	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE	942 AA;	102387 MW;	93E300B52FF549DE CRC64;
Query Match	18.8%;	Score 647.5;	DB 1; Length 942;
Best Local Similarity	30.8%;	Pred. No. 6.1e-18;	
Matches 178;	Conservative 90;	Mismatches 194;	Indels 115; Gaps 17;
QY 104	VPGPSNP-----	SREGGSPRPSPSPSPSPSDGLS-----	-----TGVVV 140
DB 426	VPGRSVNVVNINGNPDIKDKKSLSSPGSSSPSGSGSGINGDKDRGKMSSTFGIIV	485	
QY 141	GTAIGGVALLVTVTLICLLCKKRR-----	DEEDAYVYPPPPPPGKAGGPGGQQQW 195	
DB 486	GSVLGGSLSIELIGLLVFCWKYKKRQKRGSGSSNAVVHP-----	526	
QY 196	RQQATPPSDHVHTSLPPPPKAPSPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	251	
DB 527	RHSGSDNESVKTVA-----	GSSVSVGGSIDTYITLPCTSEVCE 564	
QY 252	-----PGLVLFKSKSTFTYEELARATNGFSEANLLGGGFGYVHKVLPSCKEVAVKOL	305	
DB 565	NIQMYEAGNML-----	ISIQVLRSVTNFFSSDNLILGGGFGYVHKVLPSCKEVAVKOL 618	
QY 306	KVG--SGGGEREFQAEVRLISRVHRHLVSLVGYCIAGAKRLLVTFEYFVNNLEHL---	360	
DB 619	ENGYIAGKGFARFKESEIAVLKVRHRLVTLIGYCLDNEKLVLEYMPQGTLSRHLFEW	678	
QY 361	HGEGRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIHRDIKASNLIDFKPEAKVADFGL	420	
DB 679	SEGLKPLWKQRLTALDVARVGYELHGLAQHSPIHRLKPSNILLGDMRAKVADEL 738		
QY 421	AKIASDNIHVSTRVMGTFGLAPEYAAAGKLTSEKSVFSGVLELITGRPPVDANNV	480	
DB 739	VRLAPEGKGSITRIAGTFGLAPEYAVTGRVTKVDVYSFGVILMELITGRKSLDESQP	798	
QY 481	YVEDSLVDWARPLNRASEQDFEGLADAKNNGVDEEMA--	RMVCAACVRSARR 537	
DB 799	EESHLVSWFARMK--	INKEASFKAIDITDID--LDEETLASVHTVVAELAGHCACAREPYQ 854	
QY 538	RPRMSQIVRALEGVNSLDLNEGMRPGQSNVYSYSGTSDYDSSQINDEMKKFKMALGT	597	
DB 855	RPDGHAV-----	NI-LSSLVELMKPSDONEDYIG--IDLDMSS-LPQALKAKWQAYEGRS 905	
QY 598	QEYNATG-----	EYSNPTSDYGLYPSGSSSEGG 625	
DB 906	DLESSTSLILPSLQNTQMSIPTREYGFASFTSDVGR	942	

RESULT 4

NAK_ARATH

NAK_ARATH

STANDARD: PRT; 389 AA.

P43293; Q9L296;

01-NOV-1995 (Rel. 32, Created)

15-JUN-2002 (Rel. 41, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Probable serine/threonine-protein kinase NAK [EC 2.7.1.-].

NAK Or AF5002290 OR TIE22_50.

Arabidopsis thaliana (Mouse-ear cress).

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

eurosid II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

[1]

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

MEDLINE=94032493; PubMed=8218420;

Moran T.V., Walker J.C.;

"Molecular cloning of two novel protein kinase genes from Arabidopsis thaliana";

Biochim. Biophys. Acta 1216:9-14(1993).

[2]

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

MEDLINE=21016721; PubMed=11130714;

Tabata S., Kureto T., Nakamura Y., Kotani H., Kato T., Asamizu E., Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S., Nakazaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M., Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R., Haberman J., Murray J., Johnson D., Rohlfing T., Nelson J., Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M., Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M., Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P., Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C., Wagner-McPherson C., Wollam A., Yokum M., Bell M., Dedhia N., Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J., Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M., Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I., Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T., Entlan K.-D., Terry N., Hartley N., Bent E., Johnson S., Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W., Ramsperger U., Wedler H., Balke K., Wedler E., Peters S., van Staveren M., Dirkse W., Mooijman P., Klein Lankehorst R., Teitzenger T., Bothé G., Rose M., Hauf J., Berner S., Hempel S., Feldpausch M., Lamberth S., Villarreal R., Gielen J., Ardiles W., Bents O., Lemcke K., Kolesov S., Mayer K.F.X., Rudd S., Schoof H., Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransch P.F.;

"Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana";

Nature 408:823-826(2000).

!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PLANT GROWTH AND DEVELOPMENT.

!- TISSUE SPECIFICITY: ROOTS, LEAVES AND STEMS.

!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).

EMBL; L07248; AAA18953.1; --

EMBL; AL162874; CAB85534.1; --

InterPro; IPR000719; Euk_pkinase.

InterPro; IPR004040; STY_pkinase.

InterPro; IPR002290; Ser_thr_pkinase.

Pfam; PF00069; pkinase; 1.

ProDom; PD000001; Euk_pkinase; 1.

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL: J67422; AAB09771.1; -
DR DR Maizeda; 128723; -
DR DR InterPro; IPR000719; Euk.pkinase.
DR DR InterPro; IPR004040; STY.pkinase.
DR DR InterPro; IPR002290; Ser_thr.pkinase.
DR DR InterPro; IPR001368; Tnfr_c6.
DR DR Pfam; PF00020; Tnfr_c6; 1.
DR DR Pfam; PF00069; pkinase; 1.
DR DR ProDom; PD000001; Euk.pkinase; 1.
DR DR SMART; SM00221; STYK; 1.
DR DR SMART; SM00208; TNFR; 1.
DR DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR DR PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.
DR DR PROSITE; PS00650; TNFR_NGFR_2; 1.
KW Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
FT SIGNAL 1 24
FT CHAIN 25 901
FT FT PUTATIVE RECEPTOR PROTEIN KINASE
FT CRINKL14.
FT DOMAIN 25 423
FT TRANSMEM 424 444
FT DOMAIN 445 901
FT DOMAIN 33 330
FT REPEAT 33 68
FT REPEAT 72 107
FT REPEAT 125 160
FT REPEAT 162 195
FT REPEAT 203 236
FT REPEAT 253 287
FT REPEAT 292 330
FT REPEAT 357 391
FT DOMAIN 505 712
FT NP_BIND 511 519
FT BINDING 533 533
FT ACT_SITE 634 634
FT ACT_SITE 901 AA; 97439 MW; 52F8481AC187E061 CRC64;
Query Match 17.5%; Score 605; DB 1; Length 901;
Best Local Similarity 37.2%; Pred. No. 2.3e-16;
Matches 145; Conservative 82; Mismatches 129; Indels 34; Gaps 11;
QY 263 FTYEELARATNGFSEANLLQGGFGYVHKGVLPSPGKEVAVKOLKVGSG--QGEREFQAEV 320
DB 493 FSYEELEQATGGFSDSQVQKGSFVCKILRDGTIVAVKRAIKASDVTKSKSEFNEEL 552
QY 321 EILSRVHHRLVSLGVCYIAGAKRLLYVEFPVNNLELHLEGSG---RPTMWSRLKIA 377
DB 553 DLLSLRNHALLNLLGYCEDGSESLVYEPMAHGLYQHLHGKDPNKLKRLNWARRYTIA 612
QY 378 LGSAGLSYLHEDCNKILHROIKASNILIDFEAKVADFLAKIA-SDTNTHVSTFYM 436
DB 613 VQARGIEYLHGACPPVPIHRIIDKSSNILLIDEDHNAFVADFGLSILGPADSGTPLSLPA 672
QY 437 GTFGYLAPEAASGKLTKEKSDVFSFGVVLLELTGRRPVDANNVYVDDSLVDNARPLNR 496
DB 673 GTILGLDPEYRHLTYLTKSDVYSFGVVLLEILSGKKAIDMQ--FEEGNIVEAWVPLI-- 728
QY 497 ASQGDPEGLADAKMNGYDREEMARVACAAACVRSARRRPMISQIVRALEGVSL-- 534
DB 729 --KAGDIFAILDPVSPSPDLKALKIASVACVCMRGKDRSMQKVTYTALEHALLM 786
QY 555 --SDLENGMRP-----GOSNVYSYGGSTDYDSSQZNMEDMKFKRMALGTOEYNATGEYS 607
DB 787 GSPCIEQPILPTEWILGSSRMHVKVMSQSSNHSCE-NE-----LADGEQIGG-YR 835

```

```

QY 608 NPTSDYGLYPSGSSSEGGQTREMEMGKIKR 637
DB 836 APS-WITFPSTSSQRRSSASEADIVGR 863
RESULT 6
CLVL_ARATH STANDARD; PRT: 980 AA.
ID CLVL_ARATH Q9SYQ8; O04380; Q9LQT2;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Receptor protein kinase CLAVATAL precursor (EC 2.7.1.-).
GN CLVL OR ATLG75820 OR T4012.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RX MEDLINE=97304386; PubMed=9160749;
RA Clark S.E., Williams R.W., Meyerowitz E.M.;
RT "The CLAVATAL gene encodes a putative receptor kinase that controls
RT shoot and floral meristem size in Arabidopsis.";
RL Cell 89:575-585(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99178804; PubMed=10080719;
RA Williams R.W., Clark S.E., Meyerowitz E.M.;
RT "Genetic and physical characterization of a region of Arabidopsis
RT chromosome 1 containing the CLAVATAL gene.";
RL Plant Mol. Biol. 39:171-176(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koc H.I., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
CC CC -!- FUNCTION: ACTS WITH CLV3 TO CONTROL THE BALANCE BETWEEN MERISTEM
CC CC CELL PROLIFERATION AND DIFFERENTIATION. MAY ACT WITH CLV3 AS A
CC CC LIGAND-RECEPTOR PAIR IN A SIGNAL TRANSDUCTION PATHWAY.
CC CC COORDINATING GROWTH BETWEEN ADJACENT MERISTEMATIC REGIONS.
CC CC -!- SUBUNIT: MULTIMER (POTENTIAL).
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC CC -!- TISSUE SPECIFICITY: IN A CENTRAL REGION OF THE SHOOT AND IN EARLY
CC CC FLOWER MERISTEMS.
CC CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CC -!- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way

```


Chen E., Marra M., Martienssen R., McCombie W.R.;
 "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana";
 Nature 402:769-777(1999).
 [3]
 CHARACTERIZATION.
 MEDLINE=34368830; PubMed=8086440;
 Horn M.A., Walker J.C.;
 "Biochemical properties of the autophosphorylation of RLK5, a
 receptor-like protein kinase from Arabidopsis thaliana";
 Biochim. Biophys. Acta 1208:65-74(1994).
 RL
 CC -!- COFACTOR: HAVE SIGNIFICANTLY GREATER ACTIVITY IN THE PRESENCE OF
 CC MNZ+ THAN MG2+.
 CC
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS AND ROSETTES.
 CC
 CC -!- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
 CC
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SER/THR
 CC FAMILY OF PROTEIN KINASES.
 CC
 CC -!- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
 CC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M84660; AAA32859.1; --
 CC EMBL; AL021749; CAA16865.1; --
 CC EMBL; AL161572; CAB79651.1; --
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR003592; LRR_out.
 CC InterPro: IPR002290; Ser_thr_pkinase.
 CC Pfam: PF00069; pkinase; 1.
 CC Pfam: PF00560; LRR; 20.
 CC PRINTS: PR00019; LEURICHRPT.
 CC ProDom: PD000001; Euk_pkinase; 1.
 CC SMART: SM00370; LRR; 17.
 CC
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE: PS00108; PROTEIN KINASE SE; 1.
 CC PROSITE: PS00111; PROTEIN KINASE DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
 CC Transmembrane; Glycoprotein; Phosphorylation; Leucine-rich repeat;
 CC Repeat; Signal.
 CC
 CC SIGNAL 1 14 POTENTIAL.
 CC CHAIN 15 999 RECEPTOR-LIKE PROTEIN KINASE 5.
 CC DOMAIN 15 621 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 622 641 POTENTIAL.
 CC DOMAIN 642 999 CYTOPLASMIC (POTENTIAL).
 CC REPEAT 86 112 LRR 1.
 CC REPEAT 114 138 LRR 2.
 CC REPEAT 139 161 LRR 3.
 CC REPEAT 163 186 LRR 4.
 CC REPEAT 188 211 LRR 5.
 CC REPEAT 235 261 LRR 6.
 CC REPEAT 263 283 LRR 7.
 CC REPEAT 284 306 LRR 8.
 CC REPEAT 307 330 LRR 9.
 CC REPEAT 332 353 LRR 10.
 CC REPEAT 354 378 LRR 11.
 CC REPEAT 402 426 LRR 12.
 CC REPEAT 427 450 LRR 13.
 CC REPEAT 452 474 LRR 14.
 CC REPEAT 498 522 LRR 15.
 CC REPEAT 524 547 LRR 16.
 CC REPEAT 549 567 LRR 17.
 CC REPEAT 568 592 LRR 18.
 CC DOMAIN 633 968 PROTEIN KINASE.
 CC CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC

FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NP_BIND 689 697 ATP (BY SIMILARITY).
 FT BINDING 711 711 ATP (BY SIMILARITY).
 FT ACT_SITE 819 819 BY SIMILARITY.
 FT MUTAGEN 711 711 K->E: LOSS OF CATALYTIC ACTIVITY.
 SQ SEQUENCE 999 AA; 109095 MW; F5793D899EA0C6A7 CRC64;
 Query Match 15.2%; Score 525; DB 1; Length 999;
 Best Local Similarity 36.8%; Pred. No. 2.3e-13;
 Matches 128; Conservative 61; Mismatches 121; Indels 38; Gaps 10;
 QY 258 FSKSTFTYEELARATNGFSEANLLGGGGVYVHKGLPSPGKEVAYKQLKVGSGQGERE-- 315
 DB 669 FKLHFSHEIA--DCLDEKNVIGFGSSGKVKYVELRGSEVYAVKLNKSVKGGDEYS 725
 QY 316 -----FQAEVEILSRVHRHILVINGYCIAGAKRLIVTFVFNNNLEHLHGE--GR 365
 DB 726 SDSLNRDYFAAEVETLTIRHKSIVRLWCOCSSGDKLLVTEYMPNGSLADVLHGDRKGG 785
 QY 366 PTMEWSTRLKIALGSAKLSYLHEDCNPKIHRDIKASNILEDKFEAKVADFGIAKTI-- 423
 DB 785 VVLGWPELRALDAEGLSYLHDCVPTVVEROVKSSNILLDSYDGAKVADFGIAKVGQ 845
 QY 424 -ASDINTHVSTRVMTGYLAPEYAAASKLTKESKVSFSGVVLLELITGRPPVDANNVYV 482
 DB 846 MSGSKTPEAMSGIAGSGYIAPYVYTLRVNEKSDIYSGVVVLELVTKQPTDSE--LG 903
 QY 483 DDSLVDNARPLLNRASEGDFGLADAKMNGYDREEMARVAVCAACVRSARRRRPMAS 542
 DB 904 DKMAKVVCTALDKCG-----LEPVIDPKLDLKF-KEEISKVHIHGLLCTSLPLNRP 958
 QY 543 QIVRALE---GNVSLSDLNEMGRPGQSNVYSGGSTDYDSSQXNEDM 587
 DB 959 KVVIMLQEVSGAVPCSPNTSKR-----SKTGGKL---SPYTTEDL 996
 RESULT 9
 ID IRAL_MOUSE STANDARD; PRT; 710 AA.
 AC Q62406;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Interleukin-1 receptor-associated kinase 1 (EC 2.7.1.-) (IRAK-1)
 DE (IRAK) (Pelle-like protein kinase) (mPLK).
 GN IRAK1 OR ILIRAK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RA Kopp E.B., Ghosh S.;
 FT "Cloning of mouse IRAK";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBSJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Harrington M.A.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBSJ databases.
 [3]
 RP SEQUENCE OF 34-710 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=96279287; PubMed=8663605;
 RA Trofimova M., Sprengle A.B., Green M., Sturgill T.W., Goebel M.G.,
 RA Harrington M.A.;
 FT "Developmental and tissue-specific expression of mouse pelle-like
 FT protein kinase";
 RL J. Biol. Chem. 271:17609-17612(1996).
 CC -!- FUNCTION: INVOLVED IN IL-1 PATHWAY. THIS KINASE ASSOCIATES WITH

CC THE IL-1 RECEPTOR IL1-R-1. THIS ASSOCIATION IS RAPID AND IL-1
 CC DEPENDENT (BY SIMILARITY).
 CC TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, FOLLOWED BY KIDNEY
 CC AND SKELETAL MUSCLE.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM E11 DAY TO E18 DAY.
 CC -!- PTM: AUTOPHOSPHORYLATED. AN EXTENSIVE PHOSPHORYLATION OF IRAK
 CC OCCURS AFTER ITS ASSOCIATION WITH IL1-R-1. THIS STEP COULD BE
 CC LINKED TO THE ACTIVATION OF THE KINASE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PELLE SUBFAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U56773; AAC52694.2; -
 CC EMBL; AF103876; AAD1324.1; ALT_INIT.
 CC MGD; MG1-107420; Illrak.
 CC InterPro; IPR003486; Death.
 CC InterPro; IPR003719; Euk_pkinase.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC InterPro; IPR001245; Tyr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00531; death; 1.
 CC ProDom; PD000001; Euk_pkinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC SMART; SM00219; TYRc; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 CC Transfaser; Serine/threonine-protein kinase; ATP-binding.
 CC DOMAIN 212 521 PROTEIN KINASE.
 CC FT NP_BIND 218 226 ATP (BY SIMILARITY).
 CC FT BINDING 239 239 ATP (BY SIMILARITY).
 CC FT ACT_SITE 338 338 BY SIMILARITY.
 CC SQ SEQUENCE 710 AA; 77269 MW; 8A501F002CD3EBD2 CRC64;

Query Match 14.9%; Score 515.5; DB 1; Length 710;

Best Local Similarity 26.7%; Pred. No. 3.9e-13;

Matches 174; Conservative 75; Mismatches 193; Indels 209; Gaps 25;

QY 85 TFCSPAPVTPRNPPSPVGGPPS-NPSREGG--SPRPSSSEFPP--SPSSDGLSTGVV 139

DB 100 TANHPPAVVPPS-----TAAPRPSSISAGSEAGDWSPKLQSSASTSLSPAFPGSQT--- 152

QY 140 VGIAGGVALLVITVLICLLCKKRRRDEDAYVPPPPPKAGGPGYGGQQQQWQQN 199

DB 153 -----HSELLQVLPVSLGP----- 169

QY 200 ATPPSDHWVTLPPPPKAPSPRPQPPPP-----PPPFMSGGSGSDYSDRPVLPSPS 251

DB 170 -----PLPSSADSTKSSPESVSGIQRAPSPF----- 198

QY 252 PGLVLFSEKSTFYELARATNGFSANLLQGQGGVYVHKVGLPSGKEVAVKLVKSG- 310

DB 299 -----CWPFCSISQTCNFSFELRIGEGGCGVYRAVM-RNTTYAVKRLKEADL 247

QY 311 ---QGEREFQAEVEILSRVHRLHSLVGLGYCIAGKRLIYEFVPPNNLE--LHLEGEGR 365

DB 248 ENTWVQSQELTEVEQLSRPHNPVIVFAGYCAESGLYCLVIGLPNGSLDQLHLQFQC 307

QY 366 PMWESEIRKIALGSAKGLSYLHEDCNPKIIRKASNLIDFKFEAKVADFLAKI-- 423

DB 308 SPLSPWQRDLILGTARAIQFLQHD--SPSLIHGDIKSSNVLLRLMPLKGLDGLARFSR 366

QY 424 -----ASDINTHVST-RVMGTGYLAPEVAASKLTEKSDVFGVVLLELITGRPV-- 475

DB 367 FAGAKASQSSVARTSTVGTILAYLPPEYIKTGRLAVIDTFTSGVVLLETLAGQAVRT 426

QY 476 -DANNVYVDSLVD-----WARPLLNR-----SEQGDFE 504
 DB 427 QGAKTKYLKILDEAEAEAGVTILKSTQPTLWGVATDAWAAPATAAQIYKKHLDSRPGCP 486
 QY 505 ---GLADAKNNGYDEEMARVACAAACVHSHARRPRMSQIVRALEGNVS-----LSD 556
 DB 487 POLGLA-----LAQLACC---CMHRRAKRRPPTQVYKRLGLELQAGPPMELEY 531
 QY 557 LNEGMRPGGSNYVSYGGSTDYDSSQYNEDMKFKRMALGTQE----- 599
 DB 532 AGHGSPPSQENSYMTSGAQSDEPQVPLVITRAPAAQAQQLORSPNQPVESDESYPG 591
 QY 600 YNAT-----GETSNPTSDYGLYPSGSSGQTIREMEMGKIKRTGQSY 642
 DB 592 LSATLHSHLTPGSHSPAS---PREASCTGGTTRESSV---RSSPFG 634

RESULT 10

KPRO_MAIZE

ID KPRO_MAIZE STANDARD; PRT; 817 AA.
 AC P17801;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative receptor protein kinase ZMPK1 precursor (EC 2.7.1.37).
 GN PK1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. B73; TISSUE=Root;
 RX MEDLINE=90294311; PubMed=2163028;
 RA Walker J.C. Zhang R.;
 RT "Relationship of a putative receptor protein kinase from maize to the
 RL S-locus glycoproteins of Brassica.";
 RL Nature 345:743-746(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. B73;
 RA Zhang R.; Walker J.C.;

RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PROBABLE RECEPTOR. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOPLASMIC DOMAIN.

CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE SHOOTS AND
 CC SILKS.

CC -!- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE CLOSELY RELATED TO RAF KINASES.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52384; CAA36611.1; -
 CC EMBL; X67733; CAA47962.1; -
 CC MaizeDB; 65910; -
 CC InterPro; IPR001480; B_lectin.
 CC InterPro; IPR000561; EGF-like.
 CC InterPro; IPR000719; Euk_pkinase.
 CC InterPro; IPR003014; PAN.
 CC InterPro; IPR003609; Pan_app.
 CC InterPro; IPR002290; Ser_thr_pkinase.

```

DR InterPro: IPR00058; Slocus_glycop.
DR Pfam: PF00024; ZAN; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00954; S_Locus_glycop; 1.
DR Pfam: PF01453; Agglutinin; 1.
DR ProDom: PD000001; Euk pkinase; 1.
DR SMART: SM00108; B_lectin; 1.
DR SMART: SM00181; Egf; 1.
DR SMART: SM00473; PAN_AP; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Serine/threonine-protein kinase; ATP-binding;
FT SIGNAL 1 28
FT CHAIN 29 817
FT DOMAIN 29 472 PUTATIVE RECEPTOR PROTEIN KINASE ZMPK1.
FT TRANSMEM 473 498 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 499 817 POTENTIAL.
FT DOMAIN 534 817 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 540 548 PROTEIN KINASE.
FT BINDING 562 562 ATP (BY SIMILARITY).
FT ACT_SITE 658 658 ATP (BY SIMILARITY).
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 817 AA; 91120 MW; F164B44719922E67 CRC64;

Query Match 14.8%; Score 512.5; DB 1; Length 817;
Best Local Similarity 40.1%; Pred. NO. 5.7e-13;
Matches 118; Conservative 52; Mismatches 111; Indels 13; Gaps 6;

QY 263 FYTELARATNGFSEANLGGGFGYVHKGLPSKEVAKOLKVGSGGGEREFAVEI 322
DB 524 YSYRELVKATKFKVE--LRSGETGYKGVLEDHRHVAVKLE-NVRQGEVFOAELSV 580
QY 323 ISRVHRLHSLVGYCIAGAKRLLYEFPVNNLEHLHGE-GRPTMEWSTRLKALGSA 381
DB 581 IGRINHMNLVIMGFCSEGRSLVSEVENSANILFSSGNLLDWEGRFNALGVA 640
QY 382 KGLSYLHEDCNPKLIHDIKASNTLLDPKPAKVADGLAKIAS-DNNHVSSTRVMTFG 440
DB 641 KGLAYLHHECLEWVHCDVPEPNTLLDQAFPKITDFGLVKLLNRGGSTQVSHVGRGLG 700
QY 441 YLAPEYAASGLTEKSDVSEFVGLLELITGRPPVDANNVYVDDSLVDWAPLLNRAS-- 498
DB 701 YIAPEWYSSLPITAKVDYVSGVLLLELITGRVSEL--VGGIDEVHSMKRLVRLMSAK 758
QY 499 -----EQGDFGLADAKNNNGYDREEMARMVACAAACVYHRSARRRPMSCIVRAL 548
DB 759 LEGEQSWIDGYLDSKLNRPVNYVCARTLLIKLVASCLEEDRSKRPTMEHAVQTL 812

RESULT 11
ID IRAL_HUMAN STANDARD; PRT; 712 AA.
AC P51617;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interleukin-1 receptor-associated kinase 1 (EC 2.7.1.-) (IRAK-1).
GN IRAK1 OR IRAK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96180673; PubMed=8599032;

```

```

RA RT
RL Science 271:1128-1131(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Plutzer M., Bauer D., Drescher B.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Reichwald K., Kioschis P., Rosenthal A., Plutzer M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN IL-1 PATHWAY. THIS KINASE ASSOCIATES WITH
CC THE IL-1 RECEPTOR ILI-R-1. THIS ASSOCIATION IS RAPID AND IL-1
CC DEPENDENT.
CC -!- TISSUE SPECIFICITY: SEEMS TO BE UBIQUITOUS, ALTHOUGH PRESENT IN
CC SMALL AMOUNTS.
CC -!- PTM: AN EXTENSIVE PHOSPHORYLATION OF IRAK OCCURS AFTER ITS
CC ASSOCIATION WITH ILI-R-1. THIS STEP COULD BE LINKED TO THE
CC ACTIVATION OF THE KINASE.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PELLE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; I76191; AAC41949.1; -
DR EMBL; U52112; AAC51752.1; -
DR EMBL; AF030876; AAC08756.1; -
DR EMBL; AF031075; AAF21636.1; -
DR Genew; HGNC:6112; IRAK1.
DR MIM; 300283; -
DR InterPro; IPR000488; Death.
DR InterPro; IPR000719; Euk pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00531; death; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 212 521 PROTEIN KINASE.
FT NP_BIND 218 226 ATP (BY SIMILARITY).
FT BINDING 239 239 ATP (BY SIMILARITY).
FT ACT_SITE 338 338 BY SIMILARITY.
FT CONFLICT 196 196 F -> S (IN REF. 1).
FT CONFLICT 532 532 S -> L (IN REF. 1).
SQ SEQUENCE 712 AA; 76536 MW; A7ADE07503A3961D CRC64;

Query Match 14.5%; Score 500; DB 1; Length 712;
Best Local Similarity 29.2%; Pred. NO. 1.5e-12;
Matches 173; Conservative 74; Mismatches 211; Indels 134; Gaps 24;

QY 149 LLVIVTLICLCKRRRDEEDAYVPPP-PPGPKAGPGYK-----GQQQWRQON---- 199
DB 83 LVHILTHQLL-----RADITAWHPAPLPSPGTAPRPSIIPAPAEAAWSPKLPSS 138
QY 200 -----ATPPSDHWVTSLPPPKAPSPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 252
DB 139 ASTFLSFAFCSQ--THSGDELGVPSFASLWPPPPPPSPAPSTVPG-----PESSV 187
QY 253 GLVLGFSKSTFTYE--ELARATNGFSEANLGGGFGYVHKGLPSKEVAKOLKVGSG 310
DB 188 SLGAGARPPFCWPLCEISRGTHNFSEELKIGSGFGCGVYRAVM-RNTVYAVKRLKEND 246
QY 311 ----QGEREFQAEVEIISRHHRLVSLVGYCIAGAKRLLYEFPVNNLE--LHLGEG 364
DB 247 LEWTAVKQSELTVEQLSRFRHPNIVDFAGYCAQNGFYCLVYGLPNSGLEDLHCQTQA 306

```


[illegible]

```

RESULT 15
EXTN_TOBAC
ID      EXTN_TOBAC      STANDARD;      PRT;      620 AA.
AC      P13983;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Extensin precursor (cell wall hydroxyproline-rich glycoprotein).
GN      HRGWTF3.
OS      Nicotiana tabacum (Common tobacco).

```

eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 [1]
 SEQUENCE FROM N.A.
 REP STRAIN=cv. Xanthi; TISSUE=Leaf;
 RC MEDLINE=90128263; PubMed=2612909;
 RX Kellner B., Lamb C.J.;
 RA "Specific expression of a novel cell wall hydroxyproline-rich
 RT

--!- glycoprotein gene in lateral root initiation.";
--!- RL genes Dev. 3:1639-1646(1989).
CC CC
CC CC --!- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE MAIN ROOT.
CC CC --!- SUBCELLULAR LOCATION: Extracellular matrix.
CC CC --!- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE SER-PRO(4), THE PROLINE RESIDUE IS HYDROXYLATED AND THEN GLYCOSYLATED.

CC CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

	1	20
SIGNAL	21	POTENTIAL.
CHAIN	21	EXTENSION.
REPEAT	70	H-A-P.
REPEAT	148	H-A-P-P.
REPEAT	151	H-A-P-P.

TT	DOMAIN	229	242	2	2 X 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.
TT	REPEAT	229	235	1.	1.
TT	REPEAT	229	235	2.	2.
TT	REPEAT	236	242		
TT	DOMAIN	205	620		CONTAINS THE SER-PRO(4) REPEATS.
TT	DOMAIN	499	600		3 X APPROXIMATE TANDEM REPEATS.
NO	SEQUENCE	620 AA;	65406 MW;	641DD2278AB28524	CRG64;
Query Match		10.4%;	Score 360;	DB 1;	Length 620;
Best Local Similarity		36.6%;	pred	NO. 2	1e-07.

Matches 97; Conservative 9; Mismatches 65; Indels 94; Gaps 14;

```
QY 4 APSGTSPPSPNSNTTTPPPASAPPPTTSPSSPPPTTSPSSSRSTPSAPP-S 62
Db :||| ||| ||| :||| ||| ||| ||| ||| ||| :||| |||
347 SPPPPVTSPPPPSYSP---PPPTYLPPPPSSPPPSFSP--PPPTYEQSPPPPAY 400
QY 63 PPTSTGSPPLPQSPPPATTTPGSPAPVTPTRNPPSVGPPSNPSREGGSRPPS 122
Db ||| :||| ||| ||| ||| ||| ||| ||| ||| :||| |||
401 PPLPAPPTYSPPPTYSPPPTYAOPPP---LPPTYSPPPPAYSPPPPTY---SPPPPT 454
QY 123 -SPSPSPSSDGLSTGVVVGIAIGVALLVIVILCLLCKKKRRDEEDAYVPPPP-- 179
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 YSPPP-----AYAPPPPPPT 471
QY 180 --PGKAGGYGGQQQOMROONATPPSDHVVTSLPPPKAPSPR-QPPPP-----PPPP 231
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 YSPPPPAYSP-----PPSP--IYSPPPQVQLPPTSPPPPRHLPPLPP 516
QY 232 EMSSSGSDYSDRPVLP-----PSP 252
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
517 HR-----QRPPTTYGQPPSP 533
```

Search completed: July 2, 2003, 15:22:46
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: July 2, 2003, 15:16:13 ; Search time 73 Seconds
(without alignments)
1181.002 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSAPSPGPGSPSPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3450	99.6	647	22 AAB74205	Brassica napus PER
2	2936.5	85.0	652	23 AAB92424	Herbicidally activ
3	2360.5	65.5	544	21 AAG15453	Arabidopsis thalia
4	2093.5	60.6	694	23 AAB92421	Herbicidally activ
5	1748.5	50.6	633	22 AAB74207	Protein encoded by
6	1748.5	50.6	633	23 AAB91790	Herbicidally activ
7	1729	50.1	699	23 AAB91291	Herbicidally activ
8	1687	48.9	674	22 AAB74208	Protein encoded by
9	1687	48.9	674	23 AAB93235	Herbicidally activ
10	1683.5	48.8	700	23 AAB92360	Herbicidally activ

11	1569.5	45.5	509	23 AAB92423	Herbicidally activ
12	1515.5	43.9	710	23 AAB91536	Herbicidally activ
13	1459.5	42.3	708	23 AAB91504	Herbicidally activ
14	1456	42.2	681	23 AAB93650	Herbicidally activ
15	1415.5	41.0	731	22 AAB74206	Protein encoded by
16	1400.5	40.6	308	21 AAG15454	Arabidopsis thalia
17	1399	40.5	731	22 AAB74209	Protein encoded by
18	1399	40.5	731	23 AAB93202	Herbicidally activ
19	1397	40.5	509	23 AAB91327	Herbicidally activ
20	1369.5	39.7	718	23 AAB90944	Herbicidally activ
21	1022	29.6	394	23 AAB91535	Herbicidally activ
22	816.5	23.6	707	23 AAB92807	Herbicidally activ
23	802	23.2	435	23 AAB91818	Herbicidally activ
24	791.5	22.9	753	23 AAB92299	Herbicidally activ
25	780.5	22.6	472	21 AAG31449	Arabidopsis thalia
26	780.5	22.6	472	23 AAB90797	Herbicidally activ
27	775.5	22.5	541	21 AAG45423	Arabidopsis thalia
28	774	22.4	676	23 AAB91363	Herbicidally activ
29	773	22.4	492	21 AAG31476	Arabidopsis thalia
30	773	22.4	492	23 AAB92817	Herbicidally activ
31	769.5	22.3	450	21 AAG30977	Arabidopsis thalia
32	769.5	22.3	450	23 AAB93796	Herbicidally activ
33	764	22.1	497	21 AAG46072	Herbicidally activ
34	759	22.0	492	23 AAB91378	Herbicidally activ
35	757	21.9	523	21 AAG31084	Herbicidally activ
36	757	21.9	523	23 AAB93214	Herbicidally activ
37	757	21.9	530	21 AAG36220	Arabidopsis thalia
38	757	21.9	591	23 AAO17256	A thaliana RKSIII
39	757	21.9	596	23 AAO17271	A thaliana recepto
40	757	21.9	610	21 AAG36219	Arabidopsis thalia
41	757	21.9	615	21 AAG36218	Arabidopsis thalia
42	754.5	21.9	968	21 AAB25108	Eucalyptus grandis
43	750	21.7	516	22 AAB85212	P. patens protein
44	745	21.6	1045	23 AAB91371	Herbicidally activ
45	744.5	21.6	456	21 AAG30547	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AA74205

ID AAB74205 standard; protein; 647 AA.

XX AAB74205;

XX AC

XX DT 17-MAY-2001 (first entry)

XX DE Brassica napus PERK1 protein.

XX KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant.

XX OS Brassica napus.

XX PN WO200114563-A1.

XX PD 01-MAR-2001.

XX PF 18-AUG-2000; 2000WO-CA00966.

XX PR 19-AUG-1999; 99US-0149466.

XX PR 13-OCT-1999; 99US-0159122.

XX {GORI/} GORING D.

XX {SILV/} SILVA N.

XX PI Goring D, Silva N;

XX DR WPI; 2001-244305/25.

XX PT New proline-rich, extensin-like receptor kinase nucleic acids and polypeptides useful for increasing plant wounding or pathogen

PT resistance, or for producing transgenic plants with increased wounding
 PT or pathogen resistance
 XX Claim 30; Fig 1; 91pp; English.
 XX The present invention relates to proline-rich extensin-like
 CC receptor kinase (PERK). The PERK nucleic acids and polypeptides
 CC are useful for increasing the resistance of plants to wounding
 CC and pathogens. These are also useful for producing transgenic
 CC plants with increased wounding and pathogen resistance compared
 CC with a wild type plant, as well as in assays for identifying
 CC and developing compounds to inhibit and/or enhance polypeptide
 CC function directly.
 XX Sequence 647 AA;
 SQ

Query Match 99.6%; Score 3440; DB 22; Length 647;
 Best Local Similarity 99.7%; Pred. No. 4e-181; 2; Indels 0; Gaps 0;
 Matches 645; Conservative 0; Mismatches 2;

QY 1 MSSAPSGTGGSPSPSTTTTPPPASAPPTTSSPPPTTPTSSPPSTPSAPP 60
 DB 1 MSSAPSGTGGSPSPSTTTTPPPASAPPTTSSPPPTTPTSSPPSTPSAPP 60
 QY 61 PSPTTPTSGSPPLPQSPAPPTTGGPPAVTPTTRNPPSVGPPSPNREGSGRP 120
 DB 61 PSPTTPTSGSPPLPQSPAPPTTGGPPAVTPTTRNPPSVGPPSPNREGSGRP 120
 QY 121 PSSPSPSSDGLSTGVVVGIAIGVALLVITLILCLLCKKKRRRDEDAVYVPPPP 180
 DB 121 PSSPSPSSDGLSTGVVVGIAIGVALLVITLILCLLCKKKRRRDEDAVYVPPPP 180
 QY 181 GKAGGPGGQOQWROQNAATPPSDHVTSLPPPKAPSPRPPPPPPPMSSGSD 240
 DB 181 GKAGGPGGQOQWROQNAATPPSDHVTSLPPPKAPSPRPPPPPPPMSSGSD 240
 QY 241 YSDRPVLPSPSGVLGFSKTFYEEELARATNGFSEANILGQGGFGVHGVLPSGKEV 300
 DB 241 YSDRPVLPSPSGVLGFSKTFYEEELARATNGFSEANILGQGGFGVHGVLPSGKEV 300
 QY 301 AVKOLKVGSGGGEREFOAEVEIISRVHRRHLVSLVGYCIAGAKRLLYVEEVPNNLHL 360
 DB 301 AVKOLKVGSGGGEREFOAEVEIISRVHRRHLVSLVGYCIAGAKRLLYVEEVPNNLHL 360
 QY 361 HGEGRPTMESTRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNILIDFKEAKVADFL 420
 DB 361 HGEGRPTMESTRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNILIDFKEAKVADFL 420
 QY 421 AKIASDTNTHVSTRVMTGTFGLAPEYAAAGKLTSEKSDVESGVLLELITGRRPDANNV 480
 DB 421 AKIASDTNTHVSTRVMTGTFGLAPEYAAAGKLTSEKSDVESGVLLELITGRRPDANNV 480
 QY 481 YVDSLVDMARPLINRASEGDFEGLADAKNNNGYDREEMARVACAAACVRRSARRR 540
 DB 481 YVDSLVDMARPLINRASEGDFEGLADAKNNNGYDREEMARVACAAACVRRSARRR 540
 QY 541 MSQIVRALEGNVSLDLEGMRCQSNVYSYSGSTDYDSQYNEDMKTKMALGQY 600
 DB 541 MSQIVRALEGNVSLDLEGMRCQSNVYSYSGSTDYDSQYNEDMKTKMALGQY 600
 QY 601 NATGEYSNPTSDYGLYPSGSSSGQTTREMEMKIKRTGOGYSGPSL 647
 DB 601 NATGEYSNPTSDYGLYPSGSSSGQTTREMEMKIKRTGOGYSGPSL 647

RESULT 2
 ABB92424
 ID ABB92424 standard; Protein; 652 AA.
 XX
 AC ABB92424;
 XX
 DT 31-MAY-2002 (first entry)
 XX

Herbicidally active polypeptide SEQ ID NO 1635.
 Herbicidal; plant; agriculture; herbicide.
 Arabidopsis thaliana.
 WO200210210-A2.
 07-FEB-2002.
 28-AUG-2001; 2001WO-EP09892.
 28-AUG-2001; 2001WO-EP09892.
 (FARB) BAYER AG.
 Tietjen K, Weidler M;
 WPI; 2002-269010/31.
 Identifying plant target proteins for herbicidally active compounds,
 comprising aligning and comparing nucleic acid or amino acid sequences
 from plant with nucleic acid or amino acid sequences from non-plant
 organisms -
 Claim 5; SEQ ID NO 1635; 261pp + Sequence Listing; English.
 The invention relates to identifying target proteins
 (ABB92730-ABB94016) for herbicidally active compounds, comprising
 aligning and comparing nucleic acid or amino acid sequences from plant
 with nucleic acid or amino acid sequences from non-plant organisms using
 suitable search parameters, where plant sequences having an E-value
 greater by a factor of 3 than the E-value of most similar non-plant
 sequences are selected. The polypeptides or nucleic acids encoding them
 are useful for identifying modulators. The identified modulators are
 useful as herbicides.
 Sequence 652 AA;
 Query Match 85.0%; Score 2936.5; DB 23; Length 652;
 Best Local Similarity 87.0%; Pred. No. 1.7e-153;
 Matches 574; Conservative 24; Mismatches 41; Indels 21; Gaps 12;

QY 1 MSSAPSGT--GSPSPSPSTTTTPPPA--SAPPTTSPSPPP--PSTTSPSPSSKS 54
 DB 1 MSTAPSGTTPSPSPSTTTTPPPAASSPPPTTTPSPSPSTNSTSPSPSP 60
 QY 55 TPSAPPPPTPTSPSPPLPQSPAPPTTGGPPAVTPTTRNPPSVGPPSPNPSR 113
 DB 61 PPS-PPSPSGSLT---PPLPQSPSAPITP--SPSPSTTSPNPSRSPNOGPPNTPS- 114
 QY 114 EGGSPRPSS--PSPSPSSDGLSTGVVVGIAIGVALLVITLILCLLCKKKRRR--DEE 169
 DB 115 -GSTPTPTSPNTKPSPPSPSDGLSTGVVVGIAIGVALLVITLILCLLCKKKRRRHHDE 173
 QY 170 DAYVVPPEPPPGKAGGPGGQOQWROQNAATPPSD--HVTSLPPPKAPSPRPPPP 228
 DB 174 AAYVVPPEPPPGKAGGPGGQOQWROQNAATPPSD--HVTSLPPPKAPSPRPPPP 232
 QY 229 PPP--FMSSSGSDYSDRPVLPSPSGVLGFSKTFYEEELARATNGFSEANILGQGGFG 287
 DB 233 PPPAFMSSSGSDYSDLPVLPSPSGVLGFSKTFYEEELARATNGFSEANILGQGGFG 292
 QY 288 YVHGVLPSCKEVAVKOLKVGSGGGEREFOAEVEIISRVHRRHLVSLVGYCIAGAKRLV 347
 DB 293 YVHGVLPSCKEVAVKOLKVGSGGGEREFOAEVEIISRVHRRHLVSLVGYCIAGAKRLV 352
 QY 348 YEVVNNNLEHLRGEGRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNILI 407
 DB 353 YEVVNNNLEHLRGEGRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNILI 412
 QY 408 DEKFEAKVADFLGAKIASDTNTHVSTRVMTGTFGLAPEYAAAGKLTSEKSDVESGVLLE 467

Db 413 DKFEAKVDFGLAKIASDINTHVSTRVMGTGYLAPEYAAAGKLTESDVSFGVVILE 472
QY 468 LITGRPVDANNVYDDSLVDWARPLLNRASQGDPEGLADAKMNGYDREEMARVACA 527
Db 473 LITGRPVDANNVYDDSLVDWARPLLNRASQGDPEGLADAKMNGYDREEMARVACA 532
QY 528 AACVRSARRPRMSQIVRALGNVSLDLNEGMRPGSNVTSYSGSDYDSSQYNEDM 567
Db 533 AACVRSARRPRMSQIVRALGNVSLDLNEGMRPGSNVTSYSGSDYDTSQYNDDM 592
QY 588 KFERKALQTYNQAFGEYSNPTSDYGLYPSGSSSEGTREMEMGKIKRTGGYSGPSL 647
Db 593 IKFERKALQTYNGTGEYSNPTSDYGLYPSGSSSEGTREMEMGKIKRTGGYSGPSL 652

RESULT 3
AAG15453
ID AAG15453 standard; Protein; 544 AA.
XX AC AAG15453;
DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 15711.
XX KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-030:439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123780.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132046.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142655.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

Db 348 IADGRLVYEFVFNKLEYHLGKLNLPVMEFSTRIALGAAGKGLAYLHEDCHPRIIHR 407
 QY 399 DIKASNLIDFKAEKVADEGLAKIASDINTHVSTRVMGICFYIAPAEVAAAGKTEKSDV 458
 Db 406 DIKASNLIDFNFDAWADFGLAKLTSNNTHVSTRVMTGTYGLAPAEVAAAGKTEKSDV 467
 QY 459 FSGVWLELITGRPRVDANNVYDSDIADWAPLNRASBOGDFEGLADAKMNGYDRE 518
 Db 468 FSYGVWLELITGRPRVD-NSITMDDTLVDAWAPLNRASBOGDFEGLADAKMNGYDRE 526
 QY 519 EMARWVACAAACVHRSARRPRMSQIVRALEGNVSLDNLNMGMRPGSQNVSYSGSDY 578
 Db 527 EMARWVACAAASIRHSGKRKPKMSQIVRALEGEVSLDALNMGVKGPHGSHNVYSGSDY 586
 QY 579 DSSQYNEDMKFKKMGALGTQDY---NATGEYSNPTS DYG 614
 Db 587 SQTSYNADMKKFQIALSSQEFVPSDCGEGTSSNDSRDNG 625

RESULT 6

ID ABB91790 standard; Protein; 633 AA.
 AC ABB91790;

DT 31-MAY-2002 (first entry)
 XX

DE Herbicidally active polypeptide SEQ ID NO 1001.
 XX

KW Herbicidal; plant; agriculture; herbicide.
 XX

OS Arabidopsis thaliana.
 XX

PN WO200210210-A2.
 XX

PD 07-FEB-2002.
 XX

PF 28-AUG-2001; 2001WO-EP09892.
 XX

PR 28-AUG-2001; 2001WO-EP09892.
 XX

PA (FARB) BAYER AG.
 XX

PI Tietjen K, Weidler M;
 XX

DR WPI; 2002-269010/31.
 XX

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX

PS Claim 5; SEQ ID NO 1001; 261pp + Sequence Listing; English.
 XX

CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX

SQ Sequence 633 AA;
 XX

Query Match 50.6%; Score 1748.5; DB 23; Length 633;
 Best Local Similarity 56.8%; Pred. No. 2.5e-88;
 Matches 363; Conservative 78; Mismatches 151; Indels 47; Gaps 15;

QY 3 SAPSPGTCGSPSPSSNTTTPPPASAPPTTPPSFPPPTTTPSPSSSRSTPSAPPS 62
 Db 7 SAPPTNISTSPSNTSTSSP---PAPSPFTPTPPGDSSTSSPPDSTSPAPQADN 63

QY 63 PPTSTPGSPPLPQSPAPPTTPGSPAPVTPPTNPSPSPVGPSPNSPREGSPR--P 120
 Db 64 PPN-SSNNSPSPSGGGGGERGNGNGNDTPPSRGSPS---PFSRNGDNGGSRSP 119
 QY 121 PS-----SPSPSP--SSDG-----LSTGVVWGAIGGVALLVIVLILCLCKKRR 165
 Db 120 PGDTGSRSDNPPSSGGSGGGGSRNTNTAIIVGLVAGLLMVLIVLIVLIVLIVLIV 179
 QY 166 RDEEDAYVPPPPPPGPKAGP--YGGQQQWQONATPPSDHVVVSLPPPPKAPSPR 222
 Db 180 ---DSFY-----PEPMKNGQYQYGNNNNNNA SONY--PMWHLNSQGNQXSGGWWG 227
 QY 223 QPPPPPPPPPPSSSG--GSDXS--DRPVLPPSPGLVLFKFSKSTTYELARATNGFSEA 278
 Db 228 GGPSPPPPPRMTSGEDSSMYSGPSRPVLPSPALALGFKNSTTYQELAAATGGFTDA 287
 QY 279 NLGQGGEGYVHKVLPSPGKAVAVKQLVSGSGGGERFQAEVEIISRHHRLHLSLVGYC 338
 Db 288 NLLGQGGEGYVHKVLPSPGKAVAVKSLKAGSGGGERFQAEVDIISRHHRLHLSLVGYC 347
 QY 339 IAGAKRLVYEFVFNKLEYHLGKLNLPVMEFSTRIALGAAGKGLAYLHEDCHPRIIHR 398
 Db 348 IADGRLVYEFVFNKLEYHLGKLNLPVMEFSTRIALGAAGKGLAYLHEDCHPRIIHR 407
 QY 399 DIKASNLIDFKAEKVADEGLAKIASDINTHVSTRVMTGTYGLAPAEVAAAGKTEKSDV 458
 Db 408 DIKASNLIDFNFDAWADFGLAKLTSNNTHVSTRVMTGTYGLAPAEVAAAGKTEKSDV 467
 QY 459 FSGVWLELITGRPRVDANNVYDSDIADWAPLNRASBOGDFEGLADAKMNGYDRE 518
 Db 468 FSYGVWLELITGRPRVD-NSITMDDTLVDAWAPLNRASBOGDFEGLADAKMNGYDRE 526
 QY 519 EMARWVACAAACVHRSARRPRMSQIVRALEGNVSLDNLNMGMRPGSQNVSYSGSDY 578
 Db 527 EMARWVACAAASIRHSGKRKPKMSQIVRALEGEVSLDALNMGVKGPHGSHNVYSGSDY 586
 QY 579 DSSQYNEDMKFKKMGALGTQDY---NATGEYSNPTS DYG 614
 Db 587 SQTSYNADMKKFQIALSSQEFVPSDCGEGTSSNDSRDNG 625

RESULT 7

ID ABB91291 standard; Protein; 699 AA.
 AC ABB91291;

DT 31-MAY-2002 (first entry)
 XX

DE Herbicidally active polypeptide SEQ ID NO 502.
 XX

KW Herbicidal; plant; agriculture; herbicide.
 XX

OS Arabidopsis thaliana.
 XX

PN WO200210210-A2.
 XX

PD 07-FEB-2002.
 XX

PF 28-AUG-2001; 2001WO-EP09892.
 XX

PR 28-AUG-2001; 2001WO-EP09892.
 XX

PA (FARB) BAYER AG.
 XX

PI Tietjen K, Weidler M;
 XX

DR WPI; 2002-269010/31.
 XX

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant


```

XX PN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EP09892.
XX PR 28-AUG-2001; 2001WO-EP09892.
XX PA (FARB ) BAYER AG.
XX PI Tietjen K, Weidler M;
XX PW WPI; 2002-269010/31.
XX PT Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms -
XX PS Claim 5; SEQ ID NO 715; 261pp + Sequence Listing; English.
XX XX
XX CC The invention relates to identifying target proteins
XX CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX CC aligning and comparing nucleic acid or amino acid sequences from plant
XX CC with nucleic acid or amino acid sequences from non-plant organisms using
XX CC suitable search parameters, where plant sequences having an E-value
XX CC greater by a factor of 3 than the E-value of most similar non-plant
XX CC sequences are selected. The polypeptides or nucleic acids encoding them
XX CC are useful for identifying modulators. The identified modulators are
XX CC useful as herbicides.
XX XX
XX SQ Sequence 708 AA;

Query Match 42.3%; Score 1459.5; DB 23; Length 708;
Best Local Similarity 41.8%; Pred. No. 2e-72;
Matches 317; Conservative 90; Mismatches 127; Indels 225; Gaps 21;

QY 1 MSSAP-----SGETGSPSPSSNTT-TTPPASAP-PPTIP--SSPPPTTITSP 48
DB 1 MATPTQPVNSFVTSPPPLNNAUTPATPPVTSPLPPSAPPNAPPPPPVTTSP 60
QY 49 PPSRSTP-----SAPPSPPPTSTGSPPLPQP-----SPP---APTTP 86
DB 61 PVVANGAPPPLPKPPSSPPQPVIPSPSTSPPPQVIPSPPSASPPPALVPLP 120
QY 87 GSPAP-VYTPP-----TNPPPSV-----PGPSN-PSREGGSRPPSPS 124
DB 121 SSPPPPASVPPPPSPSPPLVRSPPSVRIQSPPPPSDRPTQSPPPSPSPSERP 180
QY 125 --SPPSPSDGLSTGVVVGIAIGVALLVITLCLCKKKRDEEDAYVPPPPGP 182
DB 181 TQSPSPSPS-----ERTQSPPPPPSP-- 202
QY 183 XAGGPGYGGQQQWROONATPPSDRVVTSLEPPPK--APSPRQPPPPPPPPPPFMS----- 234
DB 203 -----SPPSDRPSQSPPPPEDTKPQPRSPNSPPTFSPPSPSP 244
QY 235 -----SSGGSDYDRPVLP-- 250
DB 245 EILVPGSNPNQNNPILRPPLDAPNTNNSGIGTGAVVGISVAVLVFTLGFIVWCLR 304
QY 251 -----SPGLV-----LGFSKST 262
DB 305 KREKRLSAYSGGDVTPSPMSSTARSDAPFRMGSAPVGVASKRSGSYQSGGIGNSKAL 364
QY 263 FTYEELARATNGFSEANLLGQGFYVHKVLPSPGKEVAVKQLKVGSGGQGERGFAEVI 322
DB 365 ESYEELVKATNGFSEANLLGEGGFGYVYKILPDGRVAVKQLKIGGQGDREFFKAET 424
QY 323 ISKVRHRLVSLVGYCIAGAKRLLYVEFVNNNLELHLHGEGRPTMEWSTRLXIALGSAK 382
DB 425 LSRIRHRLVSLVGHCHISGRDRLLIYDYSNNDLXFLHGE-KSVLDWATRVKIAGAAR 483

```

```

QY 383 GLSYLHEDCNPKLIHRDIKASNILIDFKPEAKVADFGIAKIASDINTHVSTVRMGTFGL 442
DB 484 GLAYLHEDCHPRIIHRDIKSSNILEDFNDFARVDFGLARLALDONTHTITRVIETFGYM 543
QY 443 APEVAASGKLTEKSDVRSFVGVVLLLTITGRRPVDANNVYVDDSLVDWARPLNASEQD 502
DB 544 APEYASSGKLTEKSDVRSFVGVVLLLTITGRKVPDTSQPLGDESLEWAPLISHALETEE 603
QY 503 FEGLADAKNNNGYDREEMARKVACAAACVRSARRPRMSQIVRALEGNVSLSDNEGMR 562
DB 604 FDSLADPKLGGNYVESEMFRTIEAAGACVRLHATKPRMGQIVRAFE-SLAAEDLTNGMR 662
QY 563 PQSNVSYSGSDYDSSQYNEDEMKKFKKMGALGTQYN 601
DB 663 LGESEV-----FNSAQOQSAEIRLFRMAFGSQNYS 652

```

RESULT 14

ABB93650
ID ABB93650 standard; Protein; 681 AA.

XX AC ABB93650;

XX DT 31-MAY-2002 (first entry)

XX DE Herbicidally active polypeptide SEQ ID NO 2861.

XX XX Herbicidal; plant; agriculture; herbicide.

XX OS Arabidopsis thaliana.

XX PN WO200210210-A2.

XX PD 07-FEB-2002.

XX PF 28-AUG-2001; 2001WO-EP09892.

XX PR 28-AUG-2001; 2001WO-EP09892.

XX PA (FARB) BAYER AG.

XX PI Tietjen K, Weidler M;

XX DR WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms -

XX PS Claim 5; SEQ ID NO 2861; 261pp + Sequence Listing; English.

XX CC The invention relates to identifying target proteins
XX CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX CC aligning and comparing nucleic acid or amino acid sequences from plant
XX CC with nucleic acid or amino acid sequences from non-plant organisms using
XX CC suitable search parameters, where plant sequences having an E-value
XX CC greater by a factor of 3 than the E-value of most similar non-plant
XX CC sequences are selected. The polypeptides or nucleic acids encoding them
XX CC are useful for identifying modulators. The identified modulators are
XX CC useful as herbicides.

XX SQ Sequence 681 AA;

Query Match 42.2%; Score 1456; DB 23; Length 681;

Best Local Similarity 46.4%; Pred. No. 3e-72; Indels 108; Gaps 14;

Matches 306; Conservative 87; Mismatches 158; Indels 108; Gaps 14;

QY 1 MSSAPSPGTSPPSPSPSNTTTTTPPASAPATPTTPSPSPPTTPTSPSPSRSTPSAPP 60

DB 52 VSSSPPPPVVSSP-PPSSSPSPSPVITSPPTVASSPPPVVITASPPPTPATTPAPP 110

QY 51 -----PSPPTPSTPGSP-----PPLQPPPPAPTTPGS 88
 Db 111 QTVSPPPDDASPPAPTINPKQSPSPGTPSPPTSPKPSPTPTPTTTS 170
 QY 89 PPAP-----VTPPTNP-----PPSPDGP-----PSNPSREGGSPRPPSPSP 127
 Db 171 PPPPATASPPSPNPTDPTSLAPPTPLPVYPRKPIAKPTGASNNGNLTPSS----- 226
 QY 128 SPSSDGLSTG--VVVGAIGGVALIVITLCLCKKKRRRDEEDAYVPPPPPPKAG 185
 Db 227 SPKSEVGTGIVAIGVIGVLSFVGVWFTTRKARKDGFVGTMTMP-----SAYS 282
 QY 186 GPYGGQQQWQONATPPSDHVVTSLPPPPKAPSPRQPPPPPPPPSSSGGSDYSDRP 245
 Db 283 SPQSGDVVLENSRSSAPPK-----MRSHSGSDY----- 310
 QY 246 VLPSPSGLVLGFGSKSTFTTEELARATNGSEANLGGGEGYVHKVLPSPGKEVAVKQL 305
 Db 311 MYASDSGMVSN-QRSWESYDELSQVTSFGSEKLLGEGFGCVYKGLVSDGREVAVKQL 369
 QY 305 KVGSGQGEREFQAEVEIISKVVHRLVSLVGYCIAGAKRLILVYFVPPNNLELHGEGR 365
 Db 370 KIGSGQGEREFQAEVEIISKVVHRLVSLVGYCIAGAKRLILVYFVPPNNLELHGEGR 429
 QY 366 PTMEWTRKIALGSAKGLSYLHEDCKPKIHRDIKASNILIDFKFAKVAADFGLAKIAS 425
 Db 430 PVMTWETRVRAAGAARGIAYLHEDCKHPRIHRDIKSSNILLDSFEALVADFGLAKIAQ 489
 QY 426 --DINTHVSVMGTGTYLAPEYAAACKLTKESDVTFSGVVLLELITGRPRVDANNVYD 483
 Db 490 ELDLNTHVSTRVMGTGTYLAPEYAAACKLTKESDVTFSGVVLLELITGRPRVDANNVYD 549
 QY 484 DSLVDWARPILLNRASEQDPEGLADAKMNNGYDREEMARVACAAACVRSARRPRMSQ 543
 Db 550 ESLVEMARPILLGOAIENEDEDELVDPLGNFIPGEMFVRVEAAACVRSARRPRMSQ 609
 QY 544 IVRALEGNVSLDNEGMRCQSNVYSSGGSTYDSSQYNEDMKFKRMALGQETNA 602
 Db 610 VVALDTLEATDITNGMRPGSQV-----FDSRQCSAQIRMEQMAFGSQDYSS 659

RESULT 15

AAB74206

ID AAB74206 standard; protein; 731 AA.

AC AAB74206;

XX 17-MAY-2001 (first entry)

DT Protein encoded by Arabidopsis gene #1.

DE Proline-rich extensin-like receptor kinase; PERK; resistance;

KW plant.

XW Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO200114563-A1.

XX 01-MAR-2001.

XX 18-AUG-2000; 2000WO-CA00956.

XX 19-AUG-1999; 99US-0149466.

PR 13-OCT-1999; 99US-0159122.

XX {GORI/} GORING D.

XX {SILV/} SILVA N.

XX Goring D, Silva N;

XX WPI; 2001-244305/25.

XX New proline-rich, extensin-like receptor kinase nucleic acids and

PT

polypeptides useful for increasing plant wounding or pathogen resistance, or for producing transgenic plants with increased wounding or pathogen resistance -

Examples; Fig 11; 9lpp; English.

The present invention relates to proline-rich extensin-like receptor kinase (PERK). The PERK nucleic acids and polypeptides are useful for increasing the resistance of plants to wounding and pathogens. These are also useful for producing transgenic plants with increased wounding and pathogen resistance compared with a wild type plant, as well as in assays for identifying and developing compounds to inhibit and/or enhance polypeptide function directly.

Sequence 731 AA;

Query Match 41.0%; Score 1415.5; DB 22; Length 731;
 Best Local Similarity 43.4%; Pred. No. 5.4e-70;
 Matches 321; Conservative 93; Mismatches 156; Indels 169; Gaps 22;

QY 2 SSAPSPGTGSPSPSPSNTTT-----TPPPAS-----APPTTTPSPSP 40
 Db 29 NSALPPVDSPSPSPADSSSTPLSEPSTPPDSQLPPLPSILPLTDSPPSPSPSP 88
 QY 41 PSTIPIPSPPSRSTPSAP-----PPSPPTSTFGSPSP-----LPQSPPA-----PTTP 86
 Db 89 DST-PPSPPTTSPNSPSPDESETPPAPNESNDNNPPSQDLPSPSPSPNVPNTNP 147
 QY 87 GSP--APVTPTPTNP-----PPSPVPPSPNP-----SREGGSPRPPSP--SPSP-- 129
 Db 148 ESPPLQSPAPASDPTNSPPASPLDTPNPPIQSPGATSPFANPNPSPSPFPVPPKT 207
 QY 130 -----SSDGLSTG-----VVVGIATGGVALLVIVLILCL 159
 Db 208 PSSGPVYVSPSTSPSKGTFTPNQGGDGGGGGGYQGTMTVMGAVAGFALIGVFLV 267
 QY 160 CKKKRRDERDAY-----YVPPP-----PPGPKAGGYGGQQQWROONATP 202
 Db 268 RKKAKR--NIDSYNHSQYLPHNFSVKSDGLYQDQPKGYSYSGNGSMYNNSSQSSSM 325
 QY 203 PSDHVVTSLPPPKAPSPRQPPPPPPPPPPSPSSGGSDYSDRPVLPSPGLVLGFSKST 262
 Db 326 GNSYGTAG-----GGYPHHQ-----MQSSGTPD-----SAILGSGQTH 358
 QY 263 FTYEELARATNGSEANLGGGEGYVHKVLPSPGKEVAVKOLKVGSGQGEREFQAEVEI 322
 Db 359 FSYEELAEITQGFARKNILGEGFGCVYKGTLDQKVVAVKQKAGSGQGEREFQAEVEI 418
 QY 323 ISRVHHRHLSLVGYCIAGAKRLILVYFVPPNNLELHGEGRPTMEWTRKIALGSAK 382
 Db 419 ISRVHHRHLSLVGYCISDQHRLLIEFVSQTLLEHLH-----EMSKVRITAGSAK 471
 QY 383 GLSLHEDCKPKIHRDIKASNILIDFKFAK-----VADEFLAKIA 424
 Db 472 GLAYLHEDCKPKIHRDIKASNILIDDEYEAQATMKSSFNLNLYDCKVLVADFGLARLN 531
 QY 425 SDTNTHVSTRVMGTGTYLAPEYAAACKLTKESDVTFSGVVLLELITGRPRVDANNVYD 484
 Db 532 DTTQTHVSTRVMGTGTYLAPEYAAACKLTKESDVTFSGVVLLELITGRPRVDANNVYD 591
 QY 485 SLVDWARPILLNRASEQDPEGLADAKMNNGYDREEMARVACAAACVRSARRPRMSQI 544
 Db 592 SILVENARPLLKKAETGDISELIDTRLEKRYVEHVEFRMIETAAACVRSARRPRMSQI 651
 QY 545 VRALEGNVSLDNEGMRCQSNVYSSGGSTYDSSQYNEDMKFKRMALGQETNAIG 604
 Db 652 VVALDGDGSDISNGIKIGQ-----STTYDSQYNEIDIMKFKMAFGGSDNSVESG 702
 QY 605 EYSNPTSDGLYPSGSSSE 623
 Db 703 LYS-----GNYSKSSSD 715

Search completed: July 2, 2003, 15:22:14
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 2, 2003, 14:57:12 ; Search time 18 Seconds
(without alignments)
1057.591 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTOSPPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_5/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_5/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_5/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_5/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_5/ptodata/1/iaa/PCUTUS.COMB.pep.*
6: /cgn2_5/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754.5	21.9	968	4 US-09-228-986-76	Sequence 76, Appl
2	736	21.3	630	4 US-09-228-986-71	Sequence 71, Appl
3	703.5	20.4	707	4 US-09-228-986-80	Sequence 80, Appl
4	678	19.6	638	4 US-09-228-986-74	Sequence 74, Appl
5	604.5	17.5	1196	4 US-08-881-706-2	Sequence 2, Appl
6	593	17.2	903	4 US-09-228-986-78	Sequence 78, Appl
7	579.5	16.8	980	2 US-08-473-553A-6	Sequence 6, Appl
8	575	16.7	858	2 US-08-265-628-2	Sequence 2, Appl
9	572.5	16.6	985	1 US-07-717-331F-2	Sequence 2, Appl
10	572	16.6	857	1 US-07-717-331F-2	Sequence 2, Appl
11	566.5	16.4	666	4 US-09-228-986-68	Sequence 68, Appl
12	560.5	16.2	632	4 US-09-228-986-77	Sequence 77, Appl
13	552	16.0	282	2 US-08-473-553A-4	Sequence 4, Appl
14	545	15.8	321	1 US-08-447-185-1	Sequence 1, Appl
15	545	15.8	544	2 US-08-587-680A-25	Sequence 25, Appl
16	543	15.7	655	4 US-09-228-986-70	Sequence 70, Appl
17	539.5	15.6	857	1 US-07-717-331F-3	Sequence 3, Appl
18	525	15.2	999	2 US-08-473-553A-5	Sequence 5, Appl
19	506	14.7	712	1 US-08-587-889-2	Sequence 2, Appl
20	506	14.7	712	2 US-08-980-060-5	Sequence 5, Appl
21	506	14.7	712	4 US-09-307-185-5	Sequence 5, Appl
22	506	14.7	712	5 PCI-US96-09193-2	Sequence 2, Appl
23	495	14.3	659	4 US-09-228-986-75	Sequence 75, Appl
24	486	14.1	739	4 US-09-503-922-1	Sequence 1, Appl
25	456.5	13.2	711	4 US-09-228-986-79	Sequence 79, Appl
26	430	12.5	690	4 US-09-228-986-69	Sequence 69, Appl
27	393.5	11.4	1025	2 US-08-567-375-4	Sequence 4, Appl

28	393.5	11.4	1025	2 US-08-587-680A-4	Sequence 4, Appl
29	368	10.7	1012	2 US-08-475-891A-4	Sequence 4, Appl
30	367.5	10.6	501	4 US-08-980-060-6	Sequence 6, Appl
31	367.5	10.6	501	4 US-09-307-185-6	Sequence 6, Appl
32	367.5	10.6	947	4 US-09-228-986-73	Sequence 73, Appl
33	351	10.2	1023	2 US-08-475-891A-2	Sequence 2, Appl
34	351	10.2	1023	2 US-08-567-375-2	Sequence 2, Appl
35	351	10.2	1023	2 US-08-587-680A-2	Sequence 2, Appl
36	347	10.0	913	1 US-08-445-640-4	Sequence 4, Appl
37	347	10.0	913	3 US-08-170-558-4	Sequence 4, Appl
38	347	10.0	913	3 US-08-447-314-4	Sequence 4, Appl
39	347	10.0	913	3 US-08-445-461-4	Sequence 4, Appl
40	341	9.9	919	1 US-08-336-343A-2	Sequence 2, Appl
41	331.5	9.6	330	1 US-08-642-255-32	Sequence 32, Appl
42	331.5	9.6	408	1 US-07-609-718-65	Sequence 65, Appl
43	331.5	9.6	408	1 US-08-475-411A-65	Sequence 65, Appl
44	331.5	9.6	408	4 US-08-478-029A-65	Sequence 65, Appl
45	329.5	9.5	297	2 US-08-580-545B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-228-986-76
; Sequence 76, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228, 986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-228-986-76

Query Match 21.9%; Score 754.5; DB 4; Length 968;
Best Local Similarity 35.3%; Pred. No. 1.2e-35;
Matches 184; Conservative 89; Mismatches 133; Indels 115; Gaps 16;

QY	118	PRPSSP-----SP---PSPSSDGLSTGVVIGVIGVALLVIVYLICLLCKKRRDEE	169
DB	533	PKRPFPGYFYFTASPYAFDPNGGTALSKGVIVGIAIGTFLVGLVGLYAIROKKAER	592
QY	170	DAYVYPPF-----PPPGKAGPYGGQQOQQONATPPSDHWVTSLPPPKAPSPRPP	225
DB	593	ALELSRPFASWAPSGKDSGAPOLKGRW-----	621
QY	226	PPPPPPMSSSGSDYSDRVPVLPPLPPGLVLFSGKSTFTVEELARATNGFSEANLGGG	285
DB	622	-----FSYDELKRCITNFSNDSNELGFGG	644
QY	286	FGYVHKGVLPSGKEVAVKOLKVGSGOGERFQAEVLEISRVHHHSLVLYCYAGAKRL	345
DB	645	YGVYRGLVDPGHILAIKRAQCSMQGAEFTKIELLSRVHHHKNLVLGLGFCFQEQEQM	704
QY	346	LVYEFVNNLHLHGHGGRPTMEWSTRLKIALGSAKGLSVLHDCNPKIHRDIKASNI	405
DB	705	LVYEFVNNLHLHGHGGRPTMEWSTRLKIALGSAKGLSVLHDCNPKIHRDIKASNI	764
QY	406	LIDFKEAKVADFGCLAKIASDTNT--HVSTWMTGTFGLAPEYASGKLTIEKSDVFSGV	464
DB	765	LLDEHLTAKVADFGCLAKIASDTNT--HVSTWMTGTFGLAPEYASGKLTIEKSDVFSGV	824
QY	465	LLELITGRPVDAANNVYDDSLVDWAPPLNRASEQDFEG---LADAKNN--GYDREEM	520

```

Db      825 MLESLTAKQPIEKGVVZE-----IRTDKNDQ--DIYGVREMDPMSRSMGY-LVGF 876
QY      521 ARWACAAACVRESARRRRMSQIVRALEGNVSLDNEGRMGQSNVTSYSGGSDYDS 580
Db      877 SFLDILAMCVBESADRETMSEVYKAIETMLQ-----NDGIHNSASSS---AIDFGS 929
QY      581 SQ-----YNEDMKFRKALGTQENATGEYSNPTSIDGL 615
Db      930 TKGAPRHPNDALPK-----KEVSYDSDFDYS---GGYGL 961

RESULT 2
US-09-228-986-71
; Sequence 71, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-71

Query Match      21.3%; Score 736; DB 4; Length 630;
Best Local Similarity 37.2%; Pred. No. 8.6e-35;
Matches 175; Conservative 68; Mismatches 125; Indels 102; Gaps 12;

QY      90 PAPVPTTRNPPSPVPPPP-----SNPSREGGSPRPPSPSPSPSDGLSTGVVYVIAIG 146
Db      208 PCGSPSPSPPPPPPPPPVPPVAGSNGARVQSS-----STGAIAGGYAAG 251
QY      147 VALLVITVILCLCKKRRDEEDAYVPPPPPPGPKAGPYGGQQOQROQNPATPPSDH 206
Db      252 AALLF-----AAPAIGFAW-----WRRRK---PQEH 274
QY      207 VVTSLSPPPKAPSPRPPPPPPPPPMSSSGSDYSRPPVLPSPSGVILGFSKSTFYE 266
Db      275 -----FEDVPA--EEDPEVHLGOLK--RFSLR 297
QY      267 ELARATNGFSANLLGGGFGYVHKVLPSCKEVAVKQLK-VGSGQGEREFOAEVEIISR 325
Db      298 ELQVATDGFNRNLTGRGFGKYKGRADGSLVAVKRIKEERTPGGELQFQTEVEMISM 357
QY      326 VHRHRLVSLVGYCTIAGAKRLIYVEFPVNNLE--LHLHGERPTMFWSTRUKIALGSAK 383
Db      358 AVHRLRLRLGFCMTPTERLLVVPYMWANGSVASCLRERACQNDPLDPWTRKRIALGSA 417
QY      384 LSLYLEDPCNKIIRHOIKASNILIDFKFAKVADEFLAKLASDNTHTVSTRVMTGFLA 443
Db      418 LSVLHDCDCKIIRHOVKAANILDEYEAUVGDFGLAKMDYKDTHTVAVRTGTGIIA 477
QY      444 PEYASGKLTESDVSFVGLLELTIGTRRPVANNVYDVS--LVDWARPILNRASEQ 501
Db      478 PEYLSGKSEKTDVFGYGLMELLELTIGQAFDLARLANDDDVYMLLDWVKGLL----KER 533
QY      502 DFEGLADAKNNGYDREEMARNVACAAACVYRHSARRPRMSQIVRALEGN 551
Db      534 RLDMLVDPLDKNNVVEAEVSQLTOVALICTQGSFMDPRKMSVVRMEGD 583

RESULT 3
US-09-228-986-80
; Sequence 80, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:

```

```

; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 80
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-80

Query Match      20.4%; Score 703.5; DB 4; Length 707;
Best Local Similarity 34.0%; Pred. No. 6.8e-33;
Matches 178; Conservative 81; Mismatches 179; Indels 85; Gaps 14;

QY      84 TTPGSPAPVPTTRNPPSPVPPSPNPSREGGSPRPPSPSPSPSDG-----L 134
Db      247 TGPAPPPPTTP-----PPSN-----NRPPKSSNVVPSGGSGKGNSSKSL 289
QY      135 STGVVVGIAIGVALLVITVILCLCKKRRDEEDAYVPPPPPPGPKAGPYGGQQOQ 194
Db      290 SGAIVGIIIFAVILTVAAAILGVILYARKSPREED-----EKL 329
QY      195 WQQQNA TP--PSDHVVTSLPPPKAPSPRPPPPPPPPPPPMSSSG-----GSDYSDR 247
Db      330 SNRVSTPLSLDPAELLKESPEQKVSSSPLEALAKPPPSERNKSTGDKGFSFSKRTK 389
QY      248 PPSPLVLGFSKSTFTYEELARATNGEFAENILGOGFGYVHKVLPSCKEVAVKQLK 307
Db      390 NP-----ISATYESTIADLQMA--NSFQDNLIAEGALGRYRAEFPGKILAVKLD 441
QY      308 G--SQGGEREFOAEVEIISRVRHRLVSLVGYCTIAGAKRLIYVEFPVNNLE--ELH 363
Db      442 STLSLQRPEDFLDAYSINISRLHPKTELGYCTEHEQYLLVYFDPDGLSLVDLHMADE 501
QY      364 GRPTMFWSTRUKIALGSAKGLSYLHEDCNPKIIRHDIKASNILIDFKFAKVADEFLAK 423
Db      502 TTRNLNIRKRIKIALGSAKGLSYLHEDCNPKIIRHDIKASNILIDFKFAKVADEFLAK 423
QY      424 ASDTNTHTVSTRVMTGFLAPEYASGKLTESDVSFVGLLELTIGTRRPVANNVYD 483
Db      560 ALNPNSERQVQLGSEFGISAFYVMSGIYTMKSDVYFVGMLELTIGTRKPLDSSRTSE 619
QY      484 DSLVDWARPILNRASEQDFFELA--DAKNNGYDREEMARNVACAAACVYRHSARRPR 540
Db      620 QSLVRWATPOLH-----DIDALAKMVDPAKSGYPKSLSRFADILALCIQPEFEFRPP 573
QY      541 MSQIVRALEGNVSLDNEGRMGQSNVTSYSGGSDYDSOY 583
Db      674 MSEVVQALVRMQRASLNRMTGDE-----TADHDPADY 707

RESULT 4
US-09-228-986-74
; Sequence 74, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 638
; TYPE: PRT

```


APPLICANT: Rothstein, Steven J.
 APPLICANT: Goring, Daphne
 TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
 TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESS: McAndrews, Held & Malloy, Ltd.
 STREET: 500 W. Madison St. Suite 3400
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/265,628
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/959,945
 FILING DATE:
 APPLICATION NUMBER: US 07/247,564
 FILING DATE: 03-MAR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Pochopien Ph.D., Donald J.
 REGISTRATION NUMBER: 32,167
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-707-8889
 TELEFAX: 312-707-9155
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 858 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-265-628-2

Best Local Similarity 34.8%; Pred. No. 1.1e-25;

Query Match 16.7%; Score 575; DB 2; Length 858;
Best Local Similarity 39.6%; Pred. No. 1.7e-25;
Matches 130; Conservative 65; Mismatches 120; Indels 13; Gaps 7;

[illegible]

US-08-265-628-2

; Patent No. 5821094


```

; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-473-553A-4

```

```

Query Match      16.0%; Score 552; DB 2; Length 282;
Best local Similarity 42.5%; Pred. No. 1.1e-24;
Matches 116; Conservative 53; Mismatches 94; Indels 10; Gaps 6;

QY 282 GGGGFGVHKGSLPSGKQAVKQKLVGSGGGERE--PQAEVEIISRVHHRHLSVLYGKCI 339
Db 1 GKGGSGIVYRGSPNNVDVAIKRL-VGRGTGRSDHGFATQITLGRHRRHIVRLLYGA 59

QY 340 AGAKRLVYFVFNPNLHLEHGHGRTPEWSTRLKIALGSAKGLSVLHEDCNPKIHRD 399
Db 60 NKDNTLLVEYEMNGSIGELLHSGKGLHQLQWETHRVAVZAAKGLCVLHEDCSPLIHRD 119

QY 400 IKASNIIDPKFAKQVADFGIAKIASD-INHYSTRVMGTFGLAPYAAAGSLKTEKSDV 458
Db 120 VKSNILDSDFEARVADFGIAKFLVGGASECMSSTAGSYGIAPYATLVKDEKSDV 179

QY 459 FSGVVLLELITGRPDVANNVYDLSLDVWAPL---LNRAEQGDFEGLADAKMNGY 515
Db 180 YSGVVLLELITAGKKVGFEGGVD--IVRWVNTIEETIQPSDAAIVVAIVDPRL-TGY 236

QY 515 DEEMARMVACAACVRRHSARRPRMSQIVRAL 548
Db 237 PLTSVHVFKIAMCVVEEENARPTREVVHML 269

```

```

RESULT 14
US-08-447-185-1
; Sequence 1, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman M., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1000
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-447-185-1

```

```

Query Match      15.8%; Score 545; DB 1; Length 321;
Best local Similarity 38.0%; Pred. No. 3.2e-24;
Matches 123; Conservative 57; Mismatches 132; Indels 12; Gaps 5;

QY 238 GSDYS--DRPVLPPSPGLVGLGFSKSTFYELARATNGFSEANLLSQGFGYVHKGVLP 295
Db 2 GSKYSKAINSINDALSSSYLVFFESYRVPLVDLEATNNDHFKFLIGHGVPGKVGVL 61

QY 296 SGKEVAVKQKLVGSGGGEREQAEVEIISRVHHRHLSVLYGKCIAGAKRLVYEFVNNK 355
Db 62 DGAKVALAKRRTPESGQIEEFETETETLSFCEPHPLVSLIGECDERNEMILYIKTWENGN 121

QY 356 LELHLHGGRPT--NEWSTRLKIALGSAKGLSVLHEDCNPKIHRDIKASNIIDPKFEA 413
Db 122 LKRHLVGLSDLTMSMSWQRLEICIGAARGHLHLH---TRAIHROWKSINILLDENFVP 178

QY 414 KVADFGIAKIASDIN--THVSTRVMGTFGLAPYAAAGSLKTEKSDVFGVVLLELITGR 472
Db 179 KITFGSKKTELDQTHLSTVWKTGLYIDPEYFIKGLTEKSDVYSGVYVLFVLCAR 238

QY 473 RPVDANNVYDLSLDVWAPLNLNRAEQGDFEGLADAKMNGYDREEMARMVACAACV 532
Db 239 SAIVQSLPREMNVLAEMA---VESHNGQLEQIVDPNLADKIRPESLRKFGDTAVKCLIA 294

QY 533 HSARRPRMSQIVRALESNVSLSD 556
Db 295 LSSEDPRMSGDVLWKLEYALRQ 318

```

```

RESULT 15
US-08-587-680A-25
; Sequence 25, Application US/08537680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375

```

```
/ FILING DATE: 17-JAN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/475,891
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/004,645
/ FILING DATE: 29-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/567,375
/ FILING DATE: 04-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 023070-0589400S
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 25:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 544 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..544
/ OTHER INFORMATION: /note= "Tomato Receptor Kinase 1 (TRK1)"
US-08-587-680A-25

Query Match
Best Local Similarity 15.8%; Score 545; DB 2; Length 544;
Matches 125; Conservative 59; Mismatches 111; Indels 12; Gaps 7;

QY 243 DRPVLPPSPGVLGFGSKSTTYIELARATNGFSEANLLGQGGFGYVHKGLVLPSCKEVAY 302
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 237 DNNVIGRGAGIVY---KCTMPNGDHVARSGFAAASRG----GIVYKGTMPNGDHVAV 289
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 303 KOLKVGSGQGEREFOAEVEIISRVHRLHLSLVGYCIAGAKRLLYVEFWENNLEHLHG 362
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 290 KKLGIKSGSHDGLSAELNTLGKTHRYIVRLAFCSNKEVNLVYETMLNGSLGEVLHG 349
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 363 EGRTPTMESTRLKIALGSAKGLSYLHEDCNPKTIHROIKASNILDFKFEAKVADFGLAK 422
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 350 KNGCLOWETRLKIAIEAAKGLSYLHDCSPHIIHROVKSNNILLNSELAHVADFGLAK 409
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 423 IASDTNT-HVSTRVMGTGFLAPYAAAGKLTSEKSDVSEFEGVWLELITGRPPVDANNVY 481
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 410 YFRNNGTSECMASAGSYGIAPYATLTKIDERSDYTSFGVWLELITGRPPV-GNFG 468
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 482 VDSLVDMARPLNRASEQDDEGLADAKMNGYDREEMARVACAAACVRRSARRPRM 541
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 469 EGMIVQWAK-IETKWSKEGVK-ILDERLKNVAIVEAM-QVFFVAMLCVEEYSIERPTM 525
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 542 SQIVRAL 548
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 526 REVVQML 532
Db | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
```

Search completed: July 2, 2003, 15:16:13
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 2, 2003, 15:13:41 ; Search time 35 Seconds
(without alignments)
2125.726 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGPSPSPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440963 segs, 11492915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	3453	100.0	647	9	US-10-086-464-2
2	3453	100.0	647	9	US-10-086-464-2
3	3453	100.0	721	9	US-10-086-464-5
4	1748.5	50.6	633	9	US-10-086-464-11
5	1587	48.9	674	9	US-10-086-464-14
6	1415.5	41.0	731	9	US-10-086-464-8
7	1399	40.5	731	9	US-10-086-464-17
8	754.5	21.9	963	9	US-10-101-464A-75
9	736	21.3	630	9	US-10-101-464A-71
10	722	20.9	749	10	US-09-828-313-27
11	704.5	20.4	629	9	US-10-101-464A-927
12	703.5	20.4	707	9	US-10-101-464A-80
13	678	19.6	638	9	US-10-101-464A-74
14	672	19.5	623	9	US-10-101-464A-812
15	661	19.1	366	9	US-10-219-220-263
16	660.5	19.1	515	9	US-10-101-464A-936
17	647	18.7	942	9	US-10-101-464A-911
18	641.5	18.6	894	9	US-09-754-853A-1099
19	641.5	18.6	894	9	US-09-754-853A-1116

20	641.5	18.6	894	9	US-09-754-853A-1117	Sequence 1117, Ap
21	641.5	18.6	894	9	US-09-754-853A-1118	Sequence 1118, Ap
22	641.5	18.6	894	9	US-09-754-853A-1119	Sequence 1119, Ap
23	641	18.6	744	9	US-10-101-464A-942	Sequence 942, App
24	636.5	18.4	524	9	US-10-101-464A-943	Sequence 943, App
25	635	18.4	717	9	US-10-101-464A-810	Sequence 810, App
26	626.5	18.1	365	9	US-10-101-464A-510	Sequence 510, App
27	624	18.1	919	9	US-10-101-464A-642	Sequence 642, App
28	624	18.1	1133	9	US-10-101-464A-809	Sequence 809, App
29	619	17.9	794	9	US-10-149-846-2	Sequence 2, Appli
30	612.5	17.7	502	9	US-10-101-464A-945	Sequence 945, App
31	603	17.5	901	9	US-10-208-948-16	Sequence 16, Appl
32	604.5	17.5	1021	9	US-10-101-464A-954	Sequence 954, App
33	604.5	17.5	1196	9	US-09-823-394-2	Sequence 2, Appli
34	601	17.4	355	10	US-09-862-027-26	Sequence 26, Appl
35	601	17.4	827	9	US-10-101-464A-915	Sequence 915, App
36	593	17.2	903	9	US-10-101-464A-78	Sequence 78, Appl
37	587.5	17.0	974	9	US-10-101-464A-921	Sequence 921, App
38	580.5	16.8	1014	9	US-10-101-464A-807	Sequence 807, App
39	574	16.6	339	10	US-09-838-955-8	Sequence 8, Appli
40	568.5	16.5	998	9	US-10-101-464A-895	Sequence 895, App
41	566.5	16.4	666	9	US-10-101-464A-58	Sequence 68, Appl
42	565	16.4	960	8	US-08-910-386A-18	Sequence 18, Appl
43	560.5	16.2	632	9	US-10-101-464A-77	Sequence 77, Appl
44	559	16.2	823	9	US-10-101-464A-923	Sequence 923, App
45	558	16.2	286	10	US-09-838-955-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-086-464-2
; Sequence 2, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-086-464-2

Query Match 100.0%; Score 3453; DB 9; Length 647;
Best Local Similarity 100.0%; Pred. No. 9,1e-143;
Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSAPSGTGPSPSPSPNSTTTTTPPASAPPTTTPSSPPPTTTPSSPPSRSTTPSAPP	60
Db	1	MSSAPSGTGPSPSPSPNSTTTTTPPASAPPTTTPSSPPPTTTPSSPPSRSTTPSAPP	60
QY	61	PSPTPTGSPPLPQSPAPPTGSPAPVPTTRNPPSPVPGPSKPSRSGSPRP	120
Db	61	PSPTPTGSPPLPQSPAPPTGSPAPVPTTRNPPSPVPGPSKPSRSGSPRP	120
QY	121	PSPPSPSPSSDGLSTGVVVGIAIGVALLIVITLICLCKKKRRRDEEAYVPPPPPP	180
Db	121	PSPPSPSPSSDGLSTGVVVGIAIGVALLIVITLICLCKKKRRRDEEAYVPPPPPP	180


```

QY 61 PSPT-----PSTGSPPL-----PQ-----PSPAP--TTGSPPA-PVTP-----PTN 99
Db 59 PAPPTQETSPPTSSSEVYANSPQTPENPSPPAPGSETEVTPAPPPQPSNOSP 118
QY 100 PPP-----SVPGDPSNPSR---EGGSPRPSPSPSPSSDGLST 136
Db 119 TPPSPGANDRRNRINGNNRNGSTPTSPSSNRTSGDGGSPSPRSISPQNSGDSS 178
QY 137 GWVGIAGGVALLVIVTICLLCKKKRRDEDA---YVPPPPPCPKA-GGPGGQ 191
Db 179 S---GL-----LLLLAVCICCNRAKKKSPQVNMHHYNNPNPYGGAPSGNGYK 228
QY 192 QOQTRQONATPPDHVWISLPPPKAPSPRPQPPPPPPPMSSSGSDYSDRPLPP 251
Db 229 -----TPQDHV- NMAGQGGNMGPOQPVSGP-----HSDASNLGRTAIP 271
QY 252 PGLVLGFSKSTFTYEELARATNGSEANLLQGGFGYVHKVLPSCGEVAVKQLK 311
Db 272 -AATLGHNQSTFTYDELSIATGEGFAQSNLLQGGFGYVHKVLPSCGEVAVK 330
QY 312 GEREFOAEVELISRVHRLVSLVGYCIAGAKRLVYEFVNNNLEHLHGEGRPTN 371
Db 331 GEREFOAEVDIISRVHRLVSLVGYCISGQRLVYEFVNNNLEHLHGEGRPTN 390
QY 372 TLKIALGSAKLSYLHEDC-----NPKIHRDIKASNILIDFKFEAKVAD 420
Db 391 TRVIALGSAKLSYLHEDCKKIFISHICISHPRIIHRDIKASNILIDFKFE 450
QY 421 AKIASDNTHTVSTRVMTGTFGLAPEYAAAGKLTESKSDVSEFGVLLLELIT 480
Db 451 AKLSQDNVTHVSTRVMTGTFGLAPEYASSGKLSKSDVSEFGVLLLELIT 509
QY 481 TVDSLDVWARPLNRSSEQDQFGLADAKMNGYDREEMARVACAAACVHRH 540
Db 510 EMEDSLVWARPLCLKAAQDQDYQLADPRLNLSHQMVMQASCAAALRH 569
QY 541 MSQ-----TVRALEGNVSLDLNEMRPGQSNVSYSGSTDDYSSQYNE 589
Db 570 MSQVQKLIPLVGLSILVRALEGNVSLDLNEMRPGQSNVSYSGSTDDY 629
QY 590 FRKMAIQTQYNTATGSEYNTSDYGLYPSGSSSGQITREMCMKIKRTGQ 640
Db 630 FKLALENKEYQSS-EYGG-TSEYGLNPSASSE-----EMNRGSMARN 673

```

RESULT 6

```

US-10-086-464-8
; Sequence 8, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-086-464-8
Query Match 41.0%; Score 1415.5; DB 9; Length 731;
Best Local Similarity 43.4%; Pred. No. 2e-54;

```

```

Matches 321; Conservative 93; Mismatches 156; Indels 169; Gaps 22;
QY 2 SSAPSGTGSPPSPSNSTTT-----TTPPAS-----APPTTPSPSP 40
Db 29 NSALPVDSSPPSPADSSSTPLSEPSSTPPDQSLPLPLSLPLTDSPPPSDSSPPV 88
QY 41 PSTITSPSPSPSRSTPSAP-----PPSPPTSTPGSPP-----LPQSPPA 86
Db 89 DST-PSPPPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 147
QY 87 GSPP-APVTPTRNP-----PPSPSPSPSPSP-----SRGGSPSPSPSP 129
Db 148 ESFPLQSPAPASDPTNSPPASPLDPTNPPPIQSGPATSPANANAPSPFPV 207
QY 130 -----SSDGLSTG-----VVGTAIGGVALLVITWLICLL 159
Db 208 PSSGVPVSPSLTSPSKGTPTPNQNGDGGGGGGYQKTMGMVMAVAGFAIM 267
QY 160 CKKKRRDEECAY-----YVPPP-----PPGPKAGGPGVGGQQQW 202
Db 268 RKKKR--NIDSYNHSQILPHNFVSKSDGLYQDQPGKGYSGGNGSYNNS 325
QY 203 PSDHVVTSLPPPKAPSPRPQPPPPPPPMSSSGSDYSDRPLVPPSPGLV 262
Db 326 GNSYGTAG-----GGYPHHQ-----MQSSGTPD-----SAIL 358
QY 263 FTYEELARATNGSEANLLQGGFGYVHKVLPSCGEVAVKQLKVSQGERE 322
Db 359 FTYEELAEITQGFARKNIIIGEGGFGYVKTLODGKVVAVKOLKAGSGG 418
QY 323 ISRVHRLVSLVGYCIAGAKRLVYEFVNNNLEHLHGEGRPTMENSTR 382
Db 419 ISRVHRLVSLVGYCISDOHLLIYEVSNOTLEHLH-----ENSKRV 471
QY 393 GLSYLHEDCNPKIHRDIKASNILIDFKFEAK-----VADFLGAKTA 424
Db 472 GLAYLHEDCHPKIHRDIKASNILIDDEYEAQAIMKSSFLMSLYSDCKV 531
QY 425 SDTNHVTSTRVMTGTFGLAPEYAAAGKLTESKSDVSEFGVLLLELIT 484
Db 532 DTTQTHVSTRVMTGTFGLAPEYASSGKLTDSVSEFGVLLLELITGR 591
QY 485 SLVDWARPLNRSSEQDQFGLADAKMNGYDREEMARVACAAACVHRH 544
Db 592 SLVDWARPLNRSSEQDQFGLADAKMNGYDREEMARVACAAACVHRH 651
QY 545 VRALEGNVSLDLNEMRPGQSNVSYSGSTDDYSSQYNEKMKFRMALQT 604
Db 652 VRALEGNVSLDLNEMRPGQSNVSYSGSTDDYSSQYNEKMKFRMAL 702
QY 605 EYSNPTSDYGLYPSGSSSE 623
Db 703 LYS-----GNYSKSSD 715

```

RESULT 7

```

US-10-086-464-17
; Sequence 17, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13

```



```

Query Match          20.48; Score 704.5; DB 9; Length 629;
Best Local Similarity 41.88; Pred. No. 1.1e-23;
Matches 151; Conservative 71; Mismatches 92; Indels 47; Gaps 8;

QY 263 FYVEELARATNGFSEANLGGGGYVHKGVLPSPGKEVAVKQLK---VGSQGERERQA 318
D 297 FHRELOAIANNFSSKNILGKGFGHVKGVLDQGTIVAVARLKDANAVG---GEIQPT 353
QY 319 EVELISRVHRLVSLVGYCIAGAKRLIYEFVFNNNLEHLHGEGRPTMEWSTRLKIAL 378
D 354 ELEMISLAVHRLNRLYGFCTPAEKLIIYFPMNGSVASRL--KGRPLWDGTRKIAL 411
QY 379 GSAKGLSYLHEDCNPKIIHRDIKASNLIDFKFAKVADFGAKIADTNTHVSTRVMT 438
D 412 GAAGGLYLHEDQCPKIIHROVKAAANILLDDYCEAVVGDFGLAKLHDQSHVTVAGT 471
QY 439 FGVLAPYAAAGSKLTKSDVFSFGVWLELITGRPVDD---ANNVYDDSLVDWAPLL 494
D 472 VGHIAPEYLTSGQSEKTDVFGFGLILLITGQRALEFGKAAQ---KGMILDWK--- 525
QY 495 NRASEQDDEGLADAKNNGYDREEMARVACAAACVHRSARRPRMSQIVRALEGNVSL 554
D 526 -KHOEKKLEMLVDKOLRSNYDRIELEIVQVALLCTQYLPAPHRKMSVVRMLEGD--- 581
QY 555 SDLNEGMRPGOSNYSSYGGSTDYSSOYNEDMKKFKMALGTQYXNATGEYSNPTSDYG 614
D 582 -----GLAERWASQSVESNKG-----KPEFSSSDRYSLDTPDSS 617
QY 615 L 615
D 618 L 618

RESULT 12
US-10-101-464A-80
; Sequence 80, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 1:000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-80

Query Match          20.48; Score 703.5; DB 9; Length 707;
Best Local Similarity 34.08; Pred. No. 1.4e-23;
Matches 178; Conservative 81; Mismatches 179; Indels 85; Gaps 14;

QY 84 TTPGSPAPYTPTRNPPSPVPPSPREGSPSPSPS---PPSPSDG-----L 134
D 247 TGAPPPPPPTTP-----PPSN-----NRPKSNVPPSSGSGKSNKKSL 289
QY 135 STGVVWIAIGGVALLVITLILCKKKRRRDEDAYVYPPPPPGKAGPYGGOQQ 194
D 290 SSGAIVGIIFAVILTVVAAILGVILYARKSPREDD-----BEKL 329

```

```

QY 195 WRQONATP--PSDHVVTISLPPPPKAPSPRPPPPPPPPPPPMSSSG-----GSDYSRDPVL 247
D 330 SNRVSEFTPLSPDAELLKESPEQKYSSSPLEIALKPPPSERNKSTGDKGSGISFKRTK 389
QY 248 PPPSPGLVLPFSKSTFTYEELARATNGFSEANLGGGGYVHKGVLPSPGKEVAVKQLK 307
D 390 NP-----ISATYSIADLQVATNFSQDNLIAAGALGRYRAEFPGKILAVKKLDT 441
QY 308 G--SQGGEREFOAEVEIISRVHRLVSLVGYCIAGAKRLIYEFVFNNNLEHLHGEG 363
D 442 STLSLQREDFDLDAVSNISRLHHPHITELVGYCTHEQOYLLVYFEDNSLVDLHMADE 501
QY 364 GRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNLIDFKFAKVADFGAKI 423
D 502 FTRNLSSLNRLVKIALGSAVLETLHEVCSPTVHKKFKSSNILLDDDFNPLSDCGIA-- 559
QY 424 ASDTNHSTRVMTGFGYLAPEYAAAGSKLTKSDVFSFGVWLELITGRPVDDANNVYD 483
D 550 ALNPSEERQVQVIGSGFGISAPETVMSGIYTKMSDVISFGVWLELITGRPKLDSSSTRSE 619
QY 484 DSLVDWAPLLNRASEQDDEGLA---DAKNNGYDREEMARVACAAACVHRSARRPR 540
D 620 QSLVSWATPQLH-----DIDALAKMVDPALAGSYPAKLSRFAIDIIALCQIPEPERPP 673
QY 541 MSQIVRALEGNVSLSDLNEGMRPGOSNYSSYGGSTDYSSOY 583
D 674 MSEVVOALVRMQRASLNKRMGTDE-----IADHPADY 707

RESULT 13
US-10-101-464A-74
; Sequence 74, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-74

Query Match          19.64; Score 678; DB 9; Length 638;
Best Local Similarity 43.18; Pred. No. 1.6e-22;
Matches 147; Conservative 68; Mismatches 104; Indels 22; Gaps 8;

QY 254 LVLFSGKSTFTYEELARATNGFSEANLGGGGYVHKGVLPSPGKEVAVKQLKVS-GQG 312
D 295 ICLGHKX-RFSFKELRVSTNNFNSSNIIILGVGYGIVYKGLQDGTIVAIIKRLKDNVSGG 353
QY 313 ERFQAEVEIISRVHRLVSLVGYCIAGAKRLIYEFVFNNNLE--LHLEGEGRPTMEW 370
D 354 EIQFQTEVEMISLAVHRLNRLYGFCTTSRRLRYLGYTPMNGSVASCLRDHNGKALDW 413
QY 371 STRLKIALGSAKGLSYLHEDCNPKIIHRDIKASNLIDFKFAKVADFGAKIASDTNTH 430

```

```

Db 414 PTRKRIALGAARGLLYLBHQCDPKIIHRDVKAANILLDEYFEAVVGVDFGLAKLLDHRD 473
Qy 431 VSTRVMTGFGYLAPEYASGKLTSEKSDVFSFGVWLELITGRPDV-----ANNVYVDSL 486
Db 474 VTTAVRGVGHIAPEYLSGTSGSEKTDVFGGILLLELITQRLDFGQAKQKV---M 530
Qy 487 VDWARPLNRASGDFEGLADAKMNNNGYDREEMARMVACAAACVHRSARRRPMQSIVR 546
Db 531 LDWVKKL----HQEKKLLHADKDLGNFORVELEEMVQVSLCTQFGPGRPKMCDVLR 586
Qy 547 ALEGNVSLDLNEGRPGQSNVYS-----SYGSGTDYDSS 581.
Db 587 MLEGD-GLTERWETIQIETPRYRVTELPIITYSELVEEDSS 626

RESULT 14
US-10-101-464A-812
; Sequence 812, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/101/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 1999-01-12
; PRIOR FILING DATE: 1999-01-12
; PRIOR FILING DATE: 1999-11-01
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 812
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-812

```

```

Query Match 19.5%; Score 672; DB 9; Length 623;
Best Local Similarity 40.2%; Pred. No. 2.8e-22;
Matches 144; Conservative 70; Mismatches 104; Indels 40; Gaps 7;

Qy 263 FTYELARATNGFSEANLLGGGFGYVHKVPLSGKEVAVKQLK-VGSGGGEREQAEVE 321
Db 290 FHPRELANATNFGSKNLIGRGGSHYVYRGCLSDGMVVAIKRLRNVDALGGDVQPTLR 349
Qy 322 IISRVHRHLVSLVGYCIAGAKRLLVYEFVPPNNLELHGHGEGRTMEWSTRLKALGSA 381
Db 350 MISLAVHNLRLYGFQMTQTERLLVYPYNGSVAIFL--RAKPTLEWSTKRKIALGAA 407
Qy 382 KGLSYLHEDCMPTKIIHRDIKASNLIDFKFPAKXVADFGLAKIASDTNTHVSTRVMTGFGY 441
Db 408 RGLLYLHEQCDDPKIIHRDVKAANILLDDYCEAVVDFGLAKLLDHRSHVTVAVRGTAGH 467
Qy 442 LAPEYASGKLTSEKSDVFSFGVWLELITGRPDV-----ANNVYVDSLVTDWARPLNRA 497
Db 468 IAPYLLSGSSKSTVDVFGGILLLELITGLRALFEGKAANO---RGAMLDVVR----KV 520
Qy 438 SEQGDFEGLADAKMNNNGYDREEMARMVACAAACVHRSARRRPMQSIIVRALEGNVSLDL 557
Db 521 HLEKKIEMLVDKDKONVDRIELEETIQVSLCTQYIPTNRPKMSVYVVRMLEGD----- 574
Qy 558 NEGRMPQSNVYSYSGTDYDSSQYNEDMKKFKMALGTQYENATGEYSNPTSDYGL 615
Db 575 -----GLVERKEASQRAESTR-----CRAANEFSSESYSLIDSSSL 612

```

```

RESULT 15
US-10-219-220-263
; Sequence 263, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022C1
; CURRENT APPLICATION NUMBER: US/10/219,220
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 263
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-263

Query Match 19.1%; Score 661; DB 9; Length 366;
Best Local Similarity 43.1%; Pred. No. 5.2e-22;
Matches 141; Conservative 57; Mismatches 107; Indels 22; Gaps 5;

Qy 232 FMSSSGGSDYSDRPVLPPLPPGVLGFSKSTFTYELARATNGFSEANLLGGGFGYVHK 291
Db 37 YKSEGAKEGPPQKIAPTEAPALSL-----BELKEATONFGAKALIGSGSGRVY 87
Qy 292 GVLPSGKEVAVKQLKVGSG-GGGEREQAEVEIISRVHRHLVSLVGYCIAGAKRLLVYEF 350
Db 88 AMLSQGPAAIKKLDVNSOPEANSEFLAQISMVSRLLKHDHIVELVGYCVGEGTLRLVAYEF 147
Qy 351 VPNNLELHLHG-----EGRPTMEKSTRKIALGSAKGLSYLHEDCMPTKIIHRDIKAS 403
Db 148 ATMGSLHDILHGRKGVGQAGQPGVLDVMQVRKIAVGAAGKLEYLHEKVPQPHIHRDIRS 207
Qy 404 NILIDFKFAKVAADFGLAKIASDTNTHV-STRVMTGFGYLAPEYASGKLTSEKSDVFSFG 462
Db 206 NVLLDFDFKAKIADFNLSNQAPDMAARLHSTRVLTGFGYHAPEYAMTGQIIOKSDVYSFG 267
Qy 463 VVLELITGRPREVDANNVYVDSLVTDWARPLNRASEQGDFEGLADAKMNNNGYDREEMAR 522
Db 268 VVLELITGRKPEVDHMPRGQOSLVTVATPRLS----EDKVKQCVDPKLGVEYPPKAVAK 323
Qy 523 VYACAAACVHRSARRRPMQSIIVRALE 549
Db 324 LAAVAALCVQTEADFPNMSIIVKALQ 350

Search completed: July 2, 2003, 15:20:53
Job time : 37 secs

```



```

299 EVAYKLVKVGSGQGEREQAEVEIISRYHHRHLSLVGYCI-AGAKRLLVVEFVFNNNIE 357
Db      EIAVKSLSKAGSGQGEREQAEVEIISRYHHRHLSLVGYCSNAGGQRIIVYEFNPNDIE 419
Qy      LHLHGEGRPTMEWSTRLKIALGSAGLSVLHEDCNPKIITHRIDKASNILIDFKFEAKVAD 417
Db      FLLHGKSGTYMDWTRLKIALGSAGLALYHEDCHPKIITHRIDKASNILLDNFENFAKVAD 479
Qy      FGLAKIASDTNTHVSTRVMGTFGLAPYAAAGSKLTKESDVFSFGVWILLEITGRRPVDA 477
Db      FGLAKLSQDNTHVSTRVMGTFGLAPYASSGSKLTKESDVFSFGVWILLEITGRPVDL 539
Qy      NNVTYVDSLVDWARPLNMRASEQGDFFGLADAKMNGYDREEMARWACAAACVPHSARR 537
Db      SG-DMEDSLVDWARPLCMRVADQGEYGEIVDPFFLEHGYEPEYEMARWACAAAVPHSGRR 598
Qy      RPRMSQVRALEGNVSLDLNEGMRPGCSNYSYVGGSTDYDSQYNEDMKFKRMALGT 597
Db      RPKMSQIVRTLEGASDLDLDDGYPKQSS--SGEGSSDYEMGTYGAEMKFKRVTLES 656
Qy      QEYNATGEYENPTSDYGLYPSGSSSEGGOTTREMEMG-----KIKRTGG 641
Db      RDYGASGEY-GATSEYGLDPSSSSE-----EMHIGGSTKTTTNRG 698

RESULT 3
T05264
Probable serine/threonine-specific protein kinase (EC 2.7.1.1-) T4L20.20 - Arabidopsis
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000
C:Accession: T05264
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.;
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, September 1998
A:Reference number: Z15406
A:Accession: T05264
A:Molecule type: DNA
A:Residues: 1-674 <BEV>
A:Cross-references: EMBL:AL0203094
A:Experimental source: cultivar Columbia; BAC clone T4L20
C:Genetics:
A:Map position: 4
A:Introns: 180/1; 221/1; 381/1; 421/1; 444/3; 470/2; 518/3; 563/3
A:Note: T4L20.20
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: phosphotransferase; protein kinase

```

```

Query Match      50.1%; Score 1729; DB 2; Length 699;
Best Local Similarity 52.5%; Pred. No. 1.6e-48;
Matches 372; Conservative 83; Mismatches 167; Indels 86; Gaps 19;

QY      2 SSAPSPGSGPPSPPSNS--TTTTPPA-----SAPPPTSSPPSTIPTSP--- 50
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY      9 NSSPAPPSPSPSPSSNDQTTSPFSDNQETTSFPSPSSPIAPPQQOQSESPPPLP 68
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY     51 ---SSRSTPSAPPSPPTSPTPSGSPPLLPQSPPAPT-----TP 86
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY     69 ENSDGGSSSPSPSPSSSQSQSPPP-PSTSPQOSSNNGKNNENNKNDCSSDG 127
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY     87 GSPAPVPTPTNPSPSPSVGPSPNSPRGGSPRTSPSPSPSGLSTGVVVGIAIGG 146
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    128 GNKNMSHIP----PPPSKTSDHSHSQPSRLAPPTSNSGSNSSNDGLNAGIVGLVA 183
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    147 VALLVIVTLICLLC---KKRRERDEADAYVPDPGPKAGGYGGQQQWRQNATPP 203
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    184 GILFIVMTLLVCCKFRKKKKSKLDQMYPYSNAYPAGKTGDQFYNQNAATQQOOHYNQ 243
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    204 SDHVVTSLPPE-----KAPSPPR---QPDPSPPPPFMSSSGG-----SDY 241
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    244 NDHIV-NLPPPPPGSMGTNWVSPPPPPPGNQMPSPAPV---SGGANVIQSGEMSNF 299
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    242 SD---RPVLPPSPGLVLGFKSNTFTTEELARATNGFSEANLLQGQFGYVHKVLPSCK 298
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||
QY    300 SSGPYASLPPHPESVALGPNNSTFTTEELASATQGFSDKRLLGQGGFGFYVHKGLTPNGK 359
Db       :|:| | |||||:: |||||:: |||||:: |||||:: |||||

```

```

Db 272 -AATLCHNQSTFTYDELISATEGFAQSNLLGGGFGVHKGVLPSPGKEVAVKSLKGGQ 330
QY 312 GEREFQAEVLLSRVHRHLVSLVGYCIAGAKRLLYTFVNNNLEHLHLEGRTPTWMS 371
Db 331 GEREFQAEVLLSRVHRHLVSLVGYCIAGAKRLLYTFVNNNLEHLHLEGRTPTWMS 390
QY 372 TRKIALGSAKGSYLHEDC-----NPKIHRDIKASNILLIDKFEAKVADPGL 420
Db 391 TRYKIALGSAKGSYLHEDC-----NPKIHRDIKASNILLIDKFEAKVADPGL 450
QY 421 AKTASDNTNTHSVRWMTGTFGLAPEYAAAGKLTESKDSVFGVVLLELITGRRPVANN 480
Db 451 AKLSQDNTNTHSVRWMTGTFGLAPEYAAAGKLTESKDSVFGVVLLELITGRRPVANN 509
QY 481 YVDSLDVMAKPLNCRASEGDFEGLADAKMNNGYDREEMARVACAAACVHRHARRPR 540
Db 510 EMEDSLDVMARPLCLKAAQGDYNQLADPRLNYSHOEMVOMASCAAAIRHARRRPK 569
QY 541 MSQ-----IVRLEGVNSLSDLNEMGRPGQSNVYSSYGGSTIDYSSQYNEDMKK 589
Db 570 MSQVQKLIPLVGLSILVRLEGSDMSDLSEGTREGQSTYLSPGSVSSEYDASSTADMKK 629
QY 590 FRKMALGTQYNATGEYSNPTSDYGLYPSSSSSGQTREMEMGKIKTGQ 640
Db 630 FKKLALENKEYQS-EYGG-TSEYGLNPSASSE-----EMNRGSMKRNPQ 673

RESULT 4
D96728
Hypothetical protein F24J13.3 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96728
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-710 <STO>
A:Cross-references: GB:AE005173; MID:96175155; PIDN:AAF04882.1; GSPDB:GN00141
C:Genetics:
A:Gene: F24J13.3
A:Map position: 1

```

```

Query Match 43.9%; Score 1515.5; DB 2; Length 710;
Best Local Similarity 46.2%; Pred. No. 9.3e-42;
Matches 333; Conservative 81; Mismatches 158; Indels 149; Gaps 19;

QY 2 SSAPSP-GTCSPPSPSPSNTTTPPA-----SAPPTTSPSPSPSTI 44
Db 27 SAPPTTSPSPSPSPSNTTTPPA-----SAPPTTSPSPSPSTI 83
QY 45 PTSPFPSSRTSPAPPP-----SPPTP-----STPGSP 72
Db 84 --TPPSSPPPPAPPPPIPIVFPIDSPPESTNDEPPPEVEPPPPADESPAPP 141
QY 73 PP--LPQP-----SPPATTPG--SPAPVTPPTRN-----PPPSV 104
Db 142 PPEQLPPPASSPOGPKPKKHHPGATSPAPASATSPAPPNAPPNRSHPALPKST 201
QY 105 P--GPSNPSREGSRPPSPSPSSDGLSTGVVVGIAIGVALLIVTILCLCKK 162
Db 202 AAGGLTSPSR--GVSPSGNSVPPPPANSGGYQCKTMAGFAIAGFAVIAUMAVFLVRK 259

```

```

QY 163 KRRDE--EADYVYPPP-----PPPGKAGGPGGQGGQRRQONATPPSDH 206
Db 260 KKRINDAYSCSICILPNSFNSTKSDGFLYQNPRTKISGPGGYNQSQS--NSGS 312
QY 207 VVTSLLPPPKAFSPRPQPPPPPPPMSSSGSDYSDRPVLPFPPLPPGLVGLFSKSTFTYE 266
Db 313 -----FGSQGGGGYTRSGSAPDSA--VMGSGQYHFTYE 344
QY 267 ELARATNFSANLLGQGGFGVYHKGVLPSPGKEVAVKOLKVGSGGGERFQAEVLLSRV 326
Db 345 ELTDTITEGFSKHNILGEGFGCVYAGKLDGKLVAVKOLKVGSGGQDRFKAEEVLLSRV 404
QY 327 HHRHLVSLVGYCIAGAKRLLYTFVNNNLEHLHLEGRTPTWMSNTRKIALGSAKGLSY 386
Db 405 HHRHLVSLVGYCIADSERLLYTFVNNNLEHLHLEGRTPTWMSNTRKIALGSAKGLSY 464
QY 387 LHEDCNPKIHRDIKASNILLIDKFEAKVADPGLAKIASDNTNTHSVRWMTGTFGLAPEY 446
Db 465 LHEDCNPKIHRDIKASNILLIDKFEAKVADPGLAKIASDNTNTHSVRWMTGTFGLAPEY 524
QY 447 AASGKLTESKDSVFGVVLLELITGRRPVANNVYVDSLDVMAKPLNCRASEGDFEGL 506
Db 525 AQSGLTDRSDVFGVVLLELITGRRPVANNVYVDSLDVMAKPLNCRASEGDFEGL 584
QY 507 ADAMNNGYDREEMARVACAAACVHRHARRPRMSQIVRALEGNVSLDLMGMRPGQS 566
Db 585 VDRLEKHYENEFVFMETAAACVHRHARRPRMSQIVRALEGNVSLDLMGMRPGQS 644
QY 567 NVYSSYGSTIDYSSQYNEDMKKERMALGTQYNAIGEYSNPTSDYGLYPSSSSSGQT 626
Db 645 SA-----YDSQYNNDMTKERKMAFGFDDSDSGMYS--GDYSVQDSKRGKSGNAS 692
QY 627 T 627
Db 693 S 693

RESULT 5
D96711
Hypothetical protein F24J5.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96711
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96711
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STO>
A:Cross-references: GB:AE005173; MID:95734709; PIDN:AAD49974.1; GSPDB:GN0C141
C:Genetics:
A:Gene: F24J5.8
A:Map position: 1

```

```

Query Match 42.3%; Score 1459.5; DB 2; Length 708;
Best Local Similarity 41.8%; Pred. No. 5.5e-40;
Matches 317; Conservative 90; Mismatches 127; Indels 225; Gaps 21;

QY 1 MSSAP-----SPGTGSPSPSPSPSNTTTPPPASAP--PPTTP--SSPPSPSTI 48
Db 1 MATTPVQPPVSNPSPPVTSPPPLNANSPATPPVTSPLPPSPAPPPNRPAPPPPTTSP 60
QY 49 PPSRSTP-----SAPPPSPPTPSTPGSPPLPQP-----SPP--APTTP 86

```

Db 61 PPVANGAPPPLPKPPSSPPQVIPSPPSTSPPPQVIPSPPSPASPPALVPELP 120
QY 87 GSPAP--VTPP-----TRNPPSV-----PGPSN-PSREGSPRPPSP---124
Db 121 SPPPPASVPPFRSPSPPIILVGGPPSVRFQSPPPPSDRPTQSPPPSPSPSERP 180
QY 125 --SPSSSDGLSTGVVVGIAIGVALLVITLCLCKKKRRRDEEDAYVPPPPPPGP 182
Db 181 TQSPSPSPS-----ERTQSPPPSPSP---202
QY 183 KAGPYGQQQQWQQONATPPSDHVVTLPLPPPK--APSPRQPPPPPPPPFMS-----234
Db 203 -----SPSDRPSQSPPPPPEDTKPPRRSPSPSPPTFSSPPRSP 244
QY 235 ----SSGSDSDRPVLEPP-----250
Db 245 EILVPGSNPNPQNNFTLPPDAPNPNNSGIGTGANVGVISVAVVWTLFGIVWCLR 304
QY 251 -----SPGLV-----LQFSKST 262
Db 305 KREKRLSAVSGDVTTPSPMSSTARSDGAFFRMQSSAPVGAASKRSGSYQSGLGNSKAL 364
QY 263 FTYEELARATNGFSEANLLGGGFGYVHKVLPVSGKEVAVKQKLVGSGQGEREFAVEI 322
Db 365 FSYELWKAINGFQENLLGGGFGCYKGLPDGRVAVKQKLVGSGQGEREFAVEI 424
QY 323 ISRVHHRHLSVGYCIAGAKRLIVYFVNNNLELHGHGGRPTMWSRLKLNLSAK 382
Db 425 LSRHHRHLSVGYCHCISGDRLLIYVSNNDLYHLHGB-KVLDWATKVKAAAGAR 483
QY 383 GLSYLHEDCPKTIHRDIKASNILIDFKFAKADFGAKTASDNTNTHVTRVMTGTYL 442
Db 484 GLAYLHEDCHPRIHRDIKSNILLEDNFDAVSDFGLARLALDCNTHITRVITGTYM 543
QY 443 APEYASGKLTESKSDVPSFGVWLELITGRPDVANNVYDSDLVQWAPLNRAEQGD 502
Db 544 APEYASGKLTESKSDVPSFGVWLELITGRPDVTSQPLGDESLEWAPLNRAEQGD 603
QY 503 FGLADAKMNGYDREEMARVACAAACVRSARRRPMQSIQVRALEGNVSLDLNEGMR 562
Db 604 FDSLADPPLKGNVYSESEMFRIEAAAGACVRLATKPRMQSIQVRAFE-SLAEDLTNGMR 662
QY 563 PGOSNVYSSYGSDYDSSQYNEDMKFRKALGTQYIN 601
Db 663 LGSEV-----FNSAQSAEIRLFRMAFGSQNYS 692

RESULT 6
F86387
probable Pto kinase interactor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86387
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86387
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: GB:AE005172; NID:g11079512; PIDN:AA629223.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 41.5%; Score 1434; DB 2; Length 760;
Best Local Similarity 45.9%; Pred. No. 3.7e-39;
Matches 313; Conservative 91; Mismatches 142; Indels 136; Gaps 20;

QY 3 SAPSFGTSGPPPPSPNSTTTT-----PPASAPP-----PPTPSP-PPPTSTIPT 46
Db 116 SSPPSPSPPPPPTEAPETITPISPTPTNPPPPSPSPSPSPSPSPSPSPSPSPSP 175
QY 47 S-----PPSSRSTTSAPP-----PSPPTSTP-----GSPDP-- 74
Db 176 SHSPRHLPLSPASBEIPPEPRHLPPASERSTPESDSEHSPSPSPSPSPSPSPSPSP 235
QY 75 --LPQSPAPV-----TPGSPAPV-----TPPTNPPSPVGP 108
Db 236 SKRPTSPSPSDSKRPVHPSPSPPEETLPPKSPDPLPNSNSPPTLLPSSVSP 295
QY 109 SNPSREGSPSPSPSPSP-----SSDGLSTGVVVGIAIGVALLVITL-----CLL 159
Db 296 SPKRKSVPDNP-SPNPTPTVDMSSSGISIAAVVGSIG--VALVLLTIGVWVCL 352
QY 160 CKKRRRDEEDAYVPPPPPGPKAGPGYGGQQQWQQONATPPSDHVVTLPLPPPKAPS 219
Db 353 KKKRKLSTIGGYVMTPT-----MESSPRSDSAL--LKTQSSAPL 392
QY 220 PPRQPPPPPPPFMSSSGSDYDRPVLPPSPSPGLVLPFSKSTFYEEELARATNGFSEAN 279
Db 393 V-----GNRSSNTYLSQSTFG--GFGQSRFLFYEEELVATNGFSDN 434
QY 280 LLGQGGFGYVHKVLPVSGKEVAVKQKLVGSGQGEREFAVEIISRVHHRHLSVGYCI 339
Db 435 LLGSGGFGYVHKVLPVSGKEVAVKQKLVGSGQGEREFAVEIISRVHHRHLSVGYCI 494
QY 340 AGAKRLIVYFVNNNLELHGHGGRPTMWSRLKLNLSAKGLSYLHEDCPKTIHRD 399
Db 495 SENRLLIYDVPNNLYFHLH--GTPGLDWATKVKAAAGACVRLATKPRMQSIQVRAFE 552
QY 400 IKASNILIDFKFAKADFGAKTASDNTNTHVTRVMTGTYLFGYLAPEYASGKLTESDYF 459
Db 553 IKSNILLENFHALVDFGLAKLDCNTHITRVMTGTYLFGYLAPEYASGKLTESDYF 612
QY 460 SFGVWLELITGRPDVANNVYDSDLVQWAPLNRAEQGDPLGADAKMNGYDRE 519
Db 613 SFGVWLELITGRPDVANNVYDSDLVQWAPLNRAEQGDPLGADAKMNGYDRE 579
QY 520 MARVACAAACVRSARRRPMQSIQVRALEGNVSLDLNEGMRPGQSNVYSSYGSDYD 579
Db 673 MERMIAAACIRHSATKPRMQSIQVRAFD-SLAEDLTNGMRGSEI-----IN 722
QY 580 SSQYNEDMKFRKALGTQYIN 601
Db 723 SAQSAEIRLFRMAFGSQNYS 744

RESULT 7
B86369
hypothetical protein F508.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: B86369
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alor,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzie
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86369
A:Status: preliminary
A:Molecule type: DNA

RESULT 8
T04455
hypothetical protein F4D11.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 04-Mar-2000
C:Accession: T04455
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansong, W.; Hoheisel, J.; Mewes, H.W.
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z1536Q
A:Accession: T04455
A:Molecule type: DNA

RESULT 9
A95563
probable protein kinase 60711-62822 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A95563
R:Theologits, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.-W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86239
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-715 <STO>
 A:Cross-references: GB:AE005173; NID:gi0645450; PIDN:AG21566.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F19K6.9
 A:Map position: 1

Query Match 40.5%; Score 1397; DB 2; Length 509;
 Best Local Similarity 52.5%; Pred. No. 4e-38; Indels 78; Gaps 12;
 Matches 296; Conservative 70; Mismatches 120; Indels 78; Gaps 12;
 QY 79 SPAPATPGSPAPVTPTRNPPSPVPGPPNPSREGSPSPSPSPSSDGLSTGV 138
 DB 11 SPAPATPGSPAPVTPTRNPPSPVPGPPNPSREGSPSPSPSPSSDGLSTGV 138
 QY 139 VVGTAIGCV-ALLIVITLILCKKRR--RRDEDAYVYVPPPPGKAGPYGGQQQ 194
 DB 64 ITGVVIGATFVLLGVCIFVCFYKRRKKKKKKEDIEASINRSLDPKDD---SNNLQQ 120
 QY 195 WRQCNATPPSHVWVSLPPEPKAPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 254
 DB 121 W-----SSS----- 124
 QY 255 VLGRSKTFTTEELARATNGSEANLLGQGGYGVHKGVLPSGKEVAVKQKLVGSGOGER 314
 DB 125 --EIQGNLFTYEDLSKATSNFNTNLLGQGGYGVHKGVLPSGKEVAVKQKLVGSGOGER 182
 QY 315 EFOAEVELISVHRHLVSLVGYCIAGAKLLVYEFVNNLEHLHGEGRPTMWSRL 374
 DB 183 EFOAEIQTISVHRHLVSLVGYCIAGAKLLVYEFVNNLEHLHGEGRPTMWSRL 374
 QY 375 KIALGSAGLSYLHEDCNPKLIHRDIKASNLIDFKPEAKVADFGIAKIASDTNTHVSTR 434
 DB 243 KIALGAAGLAYLHEDCNPKTIHRDVKAANLIDDSYEAKLADGLAKSSLDTHVSTR 302
 QY 435 VMGTGTYLAPYASGKLTESKDSVFSFGVILELITGRPPVDANNVYV-DPSLVDWARPL 493
 DB 303 IMGTGTYLAPYASGKLTESKDSVFSFGVILELITGRPPVDANNVYV-DPSLVDWARPL 493
 QY 494 LNRASECGDEGLADAKMNGYDREEMARVACAAACVHRSARRRPMWSQIVRALEGNVS 553
 DB 363 MIQALNKGDFGLVDPLEDFDINEMTRVACAAASVHRSARRRPMWSQIVRALEGNVS 422
 QY 554 LSDLNEMGRPGQSNVYSSYGSDTDYSSQYNEDMKFKRMALGTQYENATGEYSNPTSQY 613
 DB 423 IDDLTEGAAPGQSTIY-SLDGSSDYSTQYKEDLKPKKAFESKTFGSS-ECGSLTSDN 480
 QY 614 GLYPSGSS--EGQ-ITREMBGK 634
 DB 481 GQNPSSSSITGGR-TQIEPEK 504

RESULT 10

G86239
 protein F20824.6 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 A:Accession: G86239
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86239
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-715 <STO>
 A:Cross-references: GB:AE005172; NID:g6573752; PIDN:AAF17672.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F20B24.6
 A:Map position: 1

Query Match 35.7%; Score 1233; DB 2; Length 715;
 Best Local Similarity 39.4%; Pred. No. 8.1e-33;
 Matches 289; Conservative 84; Mismatches 163; Indels 198; Gaps 21;
 QY 2 SSAPSGTSGSPSPSNSTTTTPPPASA--PPPTSSPPPPPTPTSPSSSRSPSPAP 59
 DB 38 SETTQPPATSPSPSPDPDTQTSPPPATAAQPENQPNPTPTPTSPSSPPSLTPPPSP 97
 QY 60 PPSPTPTSTP-GSPP---PLQPSPPAP-----TT 85
 DB 98 QPQPPQSTPTGDSPPVIPPFPKQPLPPSLFPPPSLVNQLPDRPNNDNILEPINPISL 157
 QY 86 PGSPAPVTPTR-----NRP-----PSVGPSPNP--SREGGS---PAP 120
 DB 158 PFPPTPTSPSQENSGSGSPPLSLPMLPNPSEGNLPQLDPLGGSNRPVSS 217
 QY 121 PSSPSPSPSSDGLSTG-----VVGTAIGVALLVTV 153
 DB 218 SSSPSPSPSLSGNNHSGGSRNHNANGDGTSSQSNESNYTEKTVIGIGIAGVILFI 277
 QY 154 TLICLLCKKRR--RDEDAYVYVPPPPGPKAG-----GPYGGQQQOMROQNATPPSD 205
 DB 278 AGVEFVRKQKSGSSPSRSMQYLLPPAIVSVNTEGFIHYRQKPGNGSS---AQNSSPDN 334
 QY 206 HWVTSLLPPPKAPSPRQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 265
 DB 335 ----SLGNPKHGRTGPP-----DSAVIGTSKIHFTY 360
 QY 266 PELARATNGFSEANLLGQGGYGVHKGVLPSGKEVAVKQKLVGSGOGEREFAVEIISR 325
 DB 361 EELSQITGFCCKSFVVGSGGFCVYKILFEKQPAIKQKSVSAGYREFRAVEIISR 420
 QY 326 VHRHLVSLVGYCIAGAKRLVYEFVNNLEHLHGEGRPTMWSRL 364
 DB 421 VHRHLVSLVGYCIAGAKRLVYEFVNNLEHLHGEGRPTMWSRL 364
 QY 365 RPTMWSRLKALGSAKGLSYLHEDCNPKLIHRDIKASNLIDFKPEAKVADFGIAKIA 424
 DB 481 LPVLEWSRRVRIAGAAKGLAYLHEDCHPKLIHRDIKASNLIDDFEFAQ-----AK-- 532
 QY 425 SDTNTHVSTRVNGTFFGYLAPEYAAAGKLTESKDSVFSFGVILELITGRPPVDANNVYVDD 484
 DB 533 -----YLAPEYASSGKLTDRSDVFSFGVILELITGRPPVDANNVYVDD 484
 QY 485 SLVDWARPLNRAEODGDFEGDLAKMNGYDREEMARVACAAACVHRSARRRPMWSQI 544
 DB 577 SLVEWARPLLEALEKGDISEVVDPRLENDYVESEVYKNIETAAACVSHSALKRPMYQV 635
 QY 545 VRALEGNVSLDLNEMGRPGQSNVYSSYGSDTDYSSQYNEDMKFKRMA-----LGQGE 599
 DB 637 VRALTDRLDLDLNGVVKVQSRV-----YDSQYNSNEIRIFRASEDSSDLGDN- 686
 QY 600 YNATGEYSNPTSQY 613
 DB 687 ---TGYI--PSQDY 695

RESULT 11

C96728
 hypothetical protein F24J13.2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: C96728
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Jensen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M01D:21016719; PMID:11130712
A:Accession: C96728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <STO>
A:Cross-references: GB:AE005173; NID:g6175154; PIDN:AAF04881.1; GSPDB:GN00141
C:Genetics:
A:Gene: F24J13.2
A:Map position: 1
C:Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 29.6%; Score 1022; DB 2; Length 394;
Best Local Similarity 57.0%; Pred. No. 2.4e-26;
Matches 207; Conservative 54; Mismatches 82; Indels 20; Gaps 4;

QY 242 SDRPVLPPSPGL---VLGFSKSTFTYELARATNGFEANLGGGFGYVHKGLPCK 298
DB 13 ADSPSSPTAPSDSAVMGSGGTHYTYELELITGFSKQNLGGGFGCVTKGMLKDK 72
QY 299 EVAVQLKVGSGGGRFQAQVEILSRVHRLHVLGVGYCIAGAKRLVYEFVPPNNLEL 358
DB 73 LVAVQLKVGSGGGRFQAQVEILSRVHRLHVLGVGYCIADSERLLIYEVVPQTL 132
QY 359 HLHGGRPTMEWSTRLKALGSAKGLYLHEDC-----NPKIHRDITKASNLIDPKFA 413
DB 133 HLHGGRPLVEWARKVRAIVLPK---VWRICTVTSHPKLIHRDITKASNLIDDEFEV 188
QY 414 KVADFLAKIASDTHVSTRVMGTFGLAPEYASGKLTGKSDVFSFVGLLELITGR 473
DB 189 QVADFLAKVNDTTHVSTRVMGTFGLAPEYASGGLTDRSDVFSFVGLLELITGR 248
QY 474 PYDANNVYDLSVDWARPLNRAEQGDFGLADAKNNNGYDREEMARVACAAACVRRH 533
DB 249 PVDRNQPLGEESLVGARPLKKAETGDFSELVDRLKHYVKNVEPRMETAAACVRY 308
QY 534 SARRRPRMSQVLRALGNGVNSLDNEGMRPQSNVSYSGGSTDYDSOYNEDMKFKRM 593
DB 309 SGPRKPMQVLRALDSEGDMDGICNGIKVQS-----STCDDSGQNSHVTKDVGSI 360
QY 534 ALG 596
DB 361 GRG 363

RESULT 12
T01477
protein kinase homolog F1707.1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Nov-1999
C:Accession: T01477
R:Yoshitaka, Y.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li, ritz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N. submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
A:Reference number: 214334
A:Accession: T01477
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-321 <VYS>
A:Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176673; GSPDB:GN00059; ATSP:F1707

A: Experimental source: cultivar Columbia
C: Genetics:
A: Gene: ATSP.F1707.1
A: Map position: 1
A: Introns: 136/1; 166/1; 189/3; 215/2; 264/3
C: Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 28.2%; Score 973; DB 2; Length 321;
Best Local Similarity 61.7%; Pred. No. 7.le-25;
Matches 192; Conservative 45; Mismatches 62; Indels 12; Gaps 3;

QY 242 SDRPWLPPPGSL---VLGFSKSTFTVEELARATNGFSEAMLGQGFGVYHKGVLPSSK 298
Db :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
 ADSPSSPTAPSDVSANWVGSOQTHTFYEELEIDIEFGSKQNILGEGRGCVYKKLGDCK 72
QY 299 EVAVKLVKGSGQGERFPQAETIIISRVHRLHLSLVGYCIAGAKRLLVYFFVNNNLEL 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 LVAVKLVKGSGQGRDFEKAETIIISRVHRLHLSLVGYCIADSERLIYYEVNQITLKH 132
QY 359 HLHGEGRPTEWSPRLKIALGSAGLSLVLEDG-----NPKI IHRDIKASNILIDFKFEA 413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 HLGKGRPVLWARVRVIAIVLP----VVRICTKTYSHPKIIHRDINKSAILDEDEPEV 188
QY 414 KVADFGLAIASDTNTHVSRVMGTFGYLAPAYAAASKLTKEKSVTFSGVVLLELITGR 473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QVADEFELAKYNDTQTTHVSRVMGTFGYLAPAYAQSGQLTDRSDFVTFSGVVLLELITGRK 248
QY 474 PVDANNVYDDSLVDWARPLLNLRASEOGDFEGLADAKMNNGYDREEMRWACAAACVRH 533
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 PVDNRQGLGESLVGWARPILKLKAETGDFSELVDRLERHYVKNEVFMTETAACVRY 308
QY 534 SARRRPRMSQI 544
Db | :| | | | :|
 SGPKRPMMQV 319

RESULT 13
TOL1502
probable serine/threonine-specific protein kinase (EC 2.7.1.-) TLOM13.2 - Arabidopsis
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 04-Mar-2000
C: Accession: TOL1502
R: Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Martienssen, R.; McCombie, W.
submitted to the EMBL Data Library, May 1997
A: Description: The sequence of the Arabidopsis thaliana TLOM13 BAC.
A: Reference number: Z14346
A: Accession: TOL1502
A: Status: translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-707 <TOH>
A: Cross-references: EMBL:AF001308; NID:g2104523; PID:g3912917
A: Experimental source: cultivar Columbia
C: Genetics:
A: Map position: 4S
A: Introns: 132/1; 256/1; 327/1; 451/1; 532/2; 581/3
A: Note: TLOM13.2
C: Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C: Keywords: phosphotransferase

Query Match 23.6%; Score 816.5; DB 2; Length 707;
Best Local Similarity 30.2%; Pred. No. 1.2e-19;
Matches 227; Conservative 94; Mismatches 191; Indels 239; Gaps 22;

QY 10 GSPPGP---PSNSTITT--PPPASPPTTPSPPPPTIPTSPPPSSRTSPAPPSP 64
Db || | | | || | | | | | | | | | | | | | | | | | | | | | | | |
 GLPVSPTLSPSTSPVITDLPAEPFRFRKYFAPOQAEAPQSPPYSLVASDHPTSS 86
QY 65 TPSTGSPPLPQPS-----PPATTGCPSPAPVTPTNPPTS----- 103
Db || | | | || | | | | | | | | | | | | | | | | | | | | | | | |
 HFSEKSMKRNAOCPGAGLADZAPAOSNGVLDPALTQPPLSPSISNCKKSCDSMVLRKSIG 146
QY 104 -----VPGPS----- 109

Db 583 YRPKGGCUTEWARSL---EYAVEELVDPRLKRYSETQVICMHTASLCIRROPHL 638
QY 538 RPRMSQIVRALEGNYSLDLN 558
Db 639 RPRMSQVLRLEGEMLMNEIS 659

Search completed: July 2, 2003, 15:25:21
Job time : 5% secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 2, 2003, 15:19:04 ; Search time 88 Seconds
(without alignments)
1514.915 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGGPPSPNSST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_TREMBL_21.*
- 2: SP_Archaea.*
- 3: SP_Bacteria.*
- 4: SP_Fungi.*
- 5: SP_Human.*
- 6: SP_Invertebrate.*
- 7: SP_Mammal.*
- 8: SP_Mhc.*
- 9: SP_Organelle.*
- 10: SP_Phage.*
- 11: SP_Plant.*
- 12: SP_Rodent.*
- 13: SP_Virus.*
- 14: SP_Vertebrate.*
- 15: SP_Unclassified.*
- 16: SP_Rv19.*
- 17: SP_Bacteriophage.*
- 18: SP_Archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3453	100.0	647	10 Q9ARH1	Q9arhl brassica na
2	2936.5	85.0	652	10 Q9LV48	Q9lv48 arabidopsis
3	2934.5	85.0	652	10 Q9LV26	Q9lv26 arabidopsis
4	1821.5	52.8	597	10 Q8W0B8	Q8w0b8 oryza sativ
5	1769	51.2	458	10 Q9LK03	Q9lk03 arabidopsis
6	1748.5	50.6	633	10 Q9ZK08	Q9zk08 arabidopsis
7	1729	50.1	699	10 Q9X196	Q9x196 arabidopsis
8	1687	48.9	674	10 Q65672	Q65672 arabidopsis
9	1656.5	48.0	714	10 Q9LS95	Q9ls95 arabidopsis
10	1513.5	43.9	710	10 Q9CAL8	Q9cal8 arabidopsis
11	1513.5	43.8	567	10 Q9LV50	Q9lv50 arabidopsis
12	1459.5	42.3	708	10 Q9SX31	Q9sx31 arabidopsis
13	1456	42.2	681	10 Q9FTW5	Q9ftw5 arabidopsis
14	1434	41.5	760	10 Q9C660	Q9c660 arabidopsis
15	1415.5	41.0	731	10 Q9ZUE0	Q9zue0 arabidopsis
16	1399	40.5	731	10 Q65530	Q65530 arabidopsis

17	1397	40.5	509	10 Q9C821	Q9c821 arabidopsis
18	1396	40.4	698	10 Q9ASK4	Q9ask4 oryza sativ
19	1322.5	40.3	568	10 Q9AR15	Q9ar15 oryza sativ
20	1233	35.7	715	10 Q9SGY7	Q9sgy7 arabidopsis
21	1222.5	35.4	610	10 Q943G8	Q943g8 oryza sativ
22	1215	35.2	752	10 Q9LQC5	Q9lqc5 arabidopsis
23	1209.5	35.0	501	10 Q9AYA6	Q9aya6 oryza sativ
24	1103	31.9	628	10 Q9ATJ7	Q9atj7 oryza sativ
25	1022	29.6	394	10 Q9CAL9	Q9cal9 arabidopsis
26	993	28.8	378	10 Q9FW38	Q9fw38 oryza sativ
27	973	28.2	321	10 Q64591	Q64591 arabidopsis
28	927	26.8	1111	10 Q9FH74	Q9fh74 arabidopsis
29	921.5	26.7	1014	10 Q9FT38	Q9ft38 oryza sativ
30	840	24.3	744	10 Q8FW00	Q8fw00 arabidopsis
31	839	24.3	794	10 Q9SFQ2	Q9sfq2 oryza sativ
32	828.5	24.0	820	10 Q9SNF3	Q9snf3 oryza sativ
33	816.5	23.6	707	10 Q94245	Q94245 arabidopsis
34	802	23.2	435	10 Q9SK72	Q9sk72 arabidopsis
35	791.5	22.9	753	10 Q9LIC9	Q9lic9 arabidopsis
36	788.5	22.8	748	10 Q8W0N1	Q8w0n1 oryza sativ
37	787.5	22.8	479	10 Q94KD9	Q94kd9 arabidopsis
38	780.5	22.6	472	10 Q23699	Q23699 arabidopsis
39	776.5	22.5	669	10 Q93ZU1	Q93zu1 arabidopsis
40	774	22.4	676	10 Q9C890	Q9c890 arabidopsis
41	773	22.4	492	10 Q22764	Q22764 arabidopsis
42	769.5	22.3	450	10 Q04601	Q04601 arabidopsis
43	766	22.2	516	10 Q9AWQ0	Q9awq0 oryza sativ
44	764	22.1	497	10 Q9LMM7	Q9lmm7 arabidopsis
45	757	21.9	523	10 Q9SZC0	Q9szc0 arabidopsis

ALIGNMENTS

RESULT 1

Q9ARH1 ID Q9ARH1 PRELIMINARY; PRT; 647 AA.
AC Q9ARH1;
DT Q1-JUN-2001 (TREMREL. 17, Created)
DT Q1-JUN-2001 (TREMREL. 17, Last sequence update)
DT Q1-MAR-2002 (TREMREL. 20, Last annotation update)
DE Receptor protein kinase PERK1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Silva N.F., Goring D.R.;
RT "Characterization of a novel Brassica napus receptor protein kinase
RT PERK1".
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AY028699; AAK21965.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002965; P-rich_extensn.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01217; PRICHTEXTENS.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD0000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 647 AA; 68482 MW; A69AEAC0D5CA203 CRC64;

Query Match 100.0%; Score 3453; DB 10; Length 647;
Best Local Similarity 100.0%; Pred. No. 6.4e-216;

```

Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSAPSGTGGSPSPSSNTTTPPPASAPPTTPSSPPPTTPTSPSSSSSTESAPP 60
DB 1 MSSAPSGTGGSPSPSSNTTTPPPASAPPTTTPSSPPPTTPTSPSSSSSTESAPP 60
QY 61 FSPPTPTGSPPLPQSPSPAPTTPGSPAPVTPTRNPPSVFGPPSPNREGSGSPR 120
DB 61 FSPPTPTGSPPLPQSPSPAPTTPGSPAPVTPTRNPPSVFGPPSPNREGSGSPR 120
QY 121 PSSPSPSPSSDGLSTGVVGTAGVALLVITLILCLCKKRRDEDAIVYPPPPP 180
DB 121 PSSPSPSPSSDGLSTGVVGTAGVALLVITLILCLCKKRRDEDAIVYPPPPP 180
QY 181 GPKAGPYGGGQCCARQONATPPSDHVTLSLPPPKAPSPRRPQPPPPPMSSSGSD 240
DB 181 GPKAGPYGGGQCCARQONATPPSDHVTLSLPPPKAPSPRRPQPPPPPMSSSGSD 240
QY 241 YSDRVLPPSPGLVGLFSKSTFTYEELARATNGFSEANLGGOGGFGYVHKVLPGRV 300
DB 241 YSDRVLPPSPGLVGLFSKSTFTYEELARATNGFSEANLGGOGGFGYVHKVLPGRV 300
QY 301 AVKQLKVGSGGGERFQAEVETLSRVHRHLVSLVGYCIAGAKRLVVEFPNNLELHL 360
DB 301 AVKQLKVGSGGGERFQAEVETLSRVHRHLVSLVGYCIAGAKRLVVEFPNNLELHL 360
QY 361 HGEGRPTMEWTRKIALGSAKGLSYLHEDCNPKIIHRDIKASNLIDFKFEAKVADFL 420
DB 361 HGEGRPTMEWTRKIALGSAKGLSYLHEDCNPKIIHRDIKASNLIDFKFEAKVADFL 420
QY 421 AKIASDNTHTVTRVMTGFGYLAPEYASGKLTESQVSGVLELITGRPPVDANNV 480
DB 421 AKIASDNTHTVTRVMTGFGYLAPEYASGKLTESQVSGVLELITGRPPVDANNV 480
QY 481 YVDSLVDWARPLNLRASEQDFGLADAKNNNGYDREEMRVACAAACVHRSAARRPR 540
DB 481 YVDSLVDWARPLNLRASEQDFGLADAKNNNGYDREEMRVACAAACVHRSAARRPR 540
QY 541 MSQIVRALEGNVSLDNEGMKRGQSNVSYSGSTVDYSDQYEDMKFKRMALGTQY 600
DB 541 MSQIVRALEGNVSLDNEGMKRGQSNVSYSGSTVDYSDQYEDMKFKRMALGTQY 600
QY 601 NATGEYSNPTSDYGLTSGSSSEGGTREMDEMKGKTKTGQYSGPSL 647
DB 601 NATGEYSNPTSDYGLTSGSSSEGGTREMDEMKGKTKTGQYSGPSL 647

RESULT 2
Q9LV48 PRELIMINARY; PRT; 652 AA.
AC Q9LV48:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Protein kinase-like protein (A7324550/MOB24_3) (Hypothetical 59.3 kDa protein).
DE NCBI_TaxID=3702;
GN MOB24_13 OR A7324550.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko I., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT *Structural analysis of Arabidopsis thaliana chromosome 3. II.

```

```

RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones.;
RL DNA Res. 7:217-221(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT Arabidopsis cDNA clones.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB020746; BAB02007.1; -.
DR EMBL; AY059901; AAL24383.1; -.
DR EMBL; AY056788; AAL10479.1; -.
DR EMBL; AY093065; AAM13064.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002965; P-rich_extensn.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO1217; TRYCHEXTENS.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 652 AA; 69271 MW; 35005EE29FE8602F CRC64;

Query Match 85.0%; Score 2936.5; DB 10; Length 652;
Best Local Similarity 87.0%; Pred. No. 2.1e-182;
Matches 574; Conservative 24; Mismatches 41; Indels 21; Gaps 12;
QY 1 MSSAPSGT---GSPSPSPSSNTTTPPPA-SAPPTTPSPPP-PSITPTSPSSRS 54
DB 1 MSTAPSGTTPSPSPSPSTTTTPPPAASPPPTTPSPSPSPSTNSTSPSPSPL 60
QY 55 TSPAPSPPTPTGSPPLPQSPSPAPTTPGSPAPVTPTRNPPSVFGPPSPNRP 113
DB 61 PPSLPSPSPGSLT----PPLQSPSPAPITP-GPPTTSPNRPSPNQGPNTPS- 114
QY 114 EGSPSPSPSS--PSPSPSSDGLSTGVVGTAGVALLVITLILCLCKKRRR--DEE 169
DB 115 -GSTPRTPTNTKPSPPSSDGLSTGVVGTAGVALLVITLILCLCKKRRRHDE 173
QY 170 DAYTVPPPPPPKAGPYGGQCCARQONATPPSD-HVVTLSLPPPKAPSPRPPPP 229
DB 170 DAYTVPPPPPPKAGPYGGQCCARQONATPPSD-HVVTLSLPPPKAPSPRPPPP 229

```

```

Db 174 AAYVPPPPSPGKAGGPGYGGQQYKQXQNASRPSDNHVVTSL--PPKPEPPRPKPPPP 232
QY 229 PPP-FMSSSGSDYSDRVLPPSPGLWIGFSKSTFTYEELARATNGFSEANLLGQGFG 287
Db 233 PPAPFSSSGSDYSDLPVLPSPGLWIGFSKSTFTYEELSRATNGFSEANLLGQGFG 292
QY 288 YVHGVLPSGKEAVAKQLKVGSGQGEREFQAEVEIISRVHHRHLVSLVGCYIAGAKRLV 347
Db 293 YVHGILPSCGEAVAKQLKAGSGQGEREFQAEVEIISRVHHRHLVSLVGCYIAGVQRLV 352
QY 348 YEVVNNLEHLHGEGRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIHRDIKASNLI 407
Db 353 YEVVNNLEHLHGEGRPTMEWSTRLKIALGSAKGLSYLHEDCNPKIHRDIKASNLI 412
QY 408 DFKEAKVADFGGLAKIASDTNTHVSRVMTGTFGLAPEYAAAGKLTSEKSDVSEFGVLE 467
Db 413 DFKEAKVADFGGLAKIASDTNTHVSRVMTGTFGLAPEYAAAGKLTSEKSDVSEFGVLE 472
QY 468 LITGRRPVDANNVYDDSLVDWAPRLNRASEGDFGLADAKKNNGYDREEMARVACA 527
Db 473 LITGRRPVDANNVYDDSLVDWAPRLNRASEGDFGLADSKMGNEYDREEMARVACA 532
QY 528 AACVRSARRPRMSQIVRALEGNVSLDNEGMRPGQSNVYSSGSDYDSSQYNDM 587
Db 533 AACVRSARRPRMSQIVRALEGNVSLDNEGMRPGHNSVYSSGSDYDTSQYNDM 592
QY 588 KFKRMALGTQENATGEYSNPTSDYGLYPSGSSSEGTTRMEMGKIKRTGQYSGPSL 647
Db 593 IKFRKMALGTQEGYTGTEYSNPTSDYGLYPSGSSSEGTATREMENGKIKRTGQYSGPSL 652

RESULT 3
Q94JZ6 PRELIMINARY; PRT; 652 AA.
AC Q94JZ6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Protein kinase-like protein.
GN MOB24.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser I., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF370509; AAK43866.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS0108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 652 AA; 69271 MW; F2530AE5CCBD102E CRC64;

Query Match 85.0%; Score 2934.5; DB 10; Length 652;
Best Local Similarity 86.8%; Pred. No. 2.9e-182;
Matches 573; Conservative 25; Mismatches 41; Indels 21; Gaps 12;

QY 1 MSSAPSPGPT--GSPSPSPSNSTTTTPPA--SAPPTTSPSP--PSTTTPSPSSRS 54
Db 1 MSTAPSGTTPSPSPSPSTTTTPPAASPPPTTTPSPSPSPSTNSISPPSPSPL 60

```

RESULT 4

```

Q8W0B8 PRELIMINARY; PRT; 597 AA.
AC Q8W0B8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative receptor protein kinase PERK1.
GN P0452F10.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT *Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0452F10.5;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003434; BAB78668.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000847; HTH_Lysr.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

```


[illegible]

DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	ATP-binding; Hypothetical protein; Transferase.
SQ	SEQUENCE 710 AA; 75371 MW; 8B4420A8DD8F3AA6 CRC64;

Query Match	
Best Local Similarity	43.98; Score 1515.5; DB 10; Length 710;
Matches 333; Conservative	81; Mismatches 156; Indels 149; Gaps 19;

QY	2	SSAFSP--GTGSPSPSPFNSTTTTPPPA-----SAPPTTTPSSPPPPSTI 44
Db	27	SAAAPPDTSAPPSPPADS---SEPPALPSLPRAVTFSPPTVSFPPIIDSSFPDPDL 83
QY	45	PTSGPPSRSTRPSAPP-----SPPTP-----STPGSP 72
Db	84	--TPPSSPPPPADAPPPIPIVFPPPIDSPPESTNSPPPEVEFPPPPADADESPAPP 141
QY	73	PP--LPPQ-----SPAATTG-SPAATVPTRN-----PPSV 104
Db	142	PPQLPPASPDGGPKKPKKHHPGPAATSPAPATSPAPPNPAPRNSHAPKST 202
QY	105	P--GPSPNRKEGSPRPSPSPSSDGLSTGVVGIAIGGVALIVITLICLKCK 162
Db	202	AAGPLTSPR--GYSSGSNVPPPANSGGYQGKTWAGFAIAGFAIALMAVVFLVRK 259
QY	163	KRRDE--EDAYVPPP-----PPGPKAGGYGQQOQWQQOQNATPPSDH 206
Db	260	KKRIDAAYSQYLPPSFNFSTIKSDGLYGQNTKGYSGPYNSQQQS--NSGN 312
QY	207	VVISLPPPKASPSPROP PPPPPPMSSSGSDYSRVLPPSPGLVLGFSKSTTYE 265
Db	313	-----FGSQRGGGYTSGSAPOSA---VMGSGOTHTYE 344
QY	267	ELARATNGFEAMLLGGGFGYVHKVLPKGKEVAVKOLKVSGSGOGREFOAEVILSRV 326
Db	345	ELDITGEFSKHNILCEGFGCYVKLNLDGKLVAVKOLKVSGSGOGREFKAEEVILSRV 404
QY	327	HHRHLVSLGYCTAGAKRLIVFVPPNNLEHLHGGRPTMESTWRKLIALGSAKLSY 386
Db	405	HHRHLVSLGYCIADSERILLIYVNPQTLEHLHGKGRPVLEWARVRRIAISAKGLAY 464
QY	367	LHEDCNPKLIHRDIKASNILIDFFEKAVADFGLAKTASTNTHVSTRWKGTFTGYLAPEY 446
Db	465	LHEDCHPKLIHRDIKANILIDDFEAQVADFGLAKLNDSTQHVTSTRWKGTFTGYLAPEY 524
QY	447	AASGKLTESKDVSFSGVLELITGRPRVDANNVYDDSLVDWARPLLNASOGDPEGL 506
Db	525	AQSGKLTDSDFSFGVLELITGRKPVOYOPLGESLVEMARPLLHKAITGDFSEL 584
QY	507	ADAKMNNGYDREMARNVACAAVCRHSAARRPRMSQIVRALBGNTVSLDNEGMRPGQS 566
Db	585	VDRLEKKHYVENFVRMIETAAACVBSHSGPKRPMVQVVVRALDSEGDMGDISNGNKVQS 644
QY	567	NVTSYSGGSTDYDSSOYNEDMKFRKMALGTQRYNATGEYSNTSYDGLYPSSGSSSEQT 626
Db	645	SA-----YDSGGTNNDTMFKRMATCFDSSDSGMYS---GDYSVODSRKSGNEAS 692
QY	627	T 627
Db	693	S 693

RESULT 11	
ID	Q9LV50
CD	PRELIMINARY; PRT; 567 AA.
AC	Q9LV50;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Protein kinase-like protein.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
OC	Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;

```

eurosid II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Nakamura Y.;
RL MEDLINE=20363099; PubMed=10907853;
RN "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RI Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones.";
RL DNA RES. 7:217-221(2000).
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB020746; BAB02005.1; -.
DR InterPro: IPR000719; Euk.pkinase.
DR InterPro: IPR002290; Ser_thr.pkinase.
DR InterPro: IPR004040; sty.pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk.pkinase; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_S1; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 567 AA; 63160 MW; EC85F6DE26855387 CRC64;

Query Match 43.8%; Score 1513.5; DB 10; Length 567;
Best Local Similarity 64.9%; Pred. No. 3,5e-90;
Matches 311; Conservative 37; Mismatches 38; Indels 93; Gaps 7;

QY 98 RNPPSPVGPSPNPSREGSGPRPPS-----SPSPSPSSDGLSTGVTVG-AGCGVALLVI 152
DB 177 RNPP-----PPASPS--GQEPPTMTPTGFSLSPPSPR--LSTGAVVGISGGG--VFV 225
QY 153 VTILCLCKKKRRREDAYVPPPPPKAGSGYGGQQQNRQONATPPSDHVVTSLP 212
DB 226 LTLFFUCKKKRRPDDK-----ALP 245
QY 213 PPPKAPSPPPPPPPPPPPPPMSSSGSDYSDRPVLPPLPPSPGLVLGFSKSTFTYEELARAT 272
DB 246 AP-----IGLVLGHQSTFTYGLARAI 268
QY 273 NGRSEANMLQGGFGYVHKGLPSGXVAVKQLKVGSGQGERFQAEVEIISRVHHRLV 332
DB 269 NKTSEANMLGEGFGFYKIGILANGNEVAVKQLKVGSAQGEKFEQAEVNIISQIHNRNV 328
QY 333 SLVGYCTAGAKRLIVYEFVFNNNLEHLHCGEGRPTMEWSTRKLIALGSAKGLSYLHEDCN 392
DB 329 SLVGYCTAGAQRLIVYEFVFNNTLEFHLHCGKGRPTMEWSRLKIAVSSSKGLSYLHENCN 368
QY 393 PKTIHRDIKASNLIDFKFEAKYADFGLAKIASDNTHTVSTRVWGTFYGLAPYASGKL 452
DB 389 PKTIHRDIKAANILIDFKFEAKYADFGLAKIALDNTHTVSTRVWGTFYGLAPYASGKL 448
QY 453 TEKSDVSEGCWLELITGRPRPDANNVYDDSLVDWARPLLNKASEQDGFGLADAKMN 512
DB 449 TEKSDVYSGVWLELITGRPPVDANNVYADDSLVDWARPLVQALLESNFGLADIKN 508
QY 513 NGVDREMARVACAACVRRHSARRPRMSQIVRALEGNVISDLNEMGRMPGOSWYSS 571
DB 509 NEVDREMARVACAACVRYTARRPRMDQVVRVLEGNISPSDLAQGITPGHSWYSS 567

RESULT 12
Q9SX31 PRELIMINARY; PRT; 708 AA.
AC Q9SX31;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```

```

01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE F24J5.8 protein (Hypothetical 75.1 kDa protein).
GN F24J5.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,
RA Chin C., Choi E., Chiou J., Altafi H., Araujo R., Brooks S.,
RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
RA Hwang B., Huizar L., Khan S., Kim C., Palm C., Rowley D., Shinn P.,
RA Walker M., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.,
RT "The sequence of BAC F24J5 from Arabidopsis thaliana chromosome 1.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
RT "Full length cDNA of gene F24J5.8 (GI:5734709).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AC008075; AAD49974.1; -.
DR EMBL: AY035076; AAK59581.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002965; P-rich_extensions.
DR InterPro: IPR004290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR01217; PRICHEXTENS.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00221; STYK.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
KW Transferase.
SQ
SEQUENCE 708 AA; 75127 MW; AC0818E54B076328 CRC64;
Query Match 42.3%; Score 1459.5; DB 10; Length 708;
Best Local Similarity 41.8%; Pred. No. 1.5e-86;
Matches 317; Conservative 90; Mismatches 127; Indels 225; Gaps 21;
QY 1 MCSAP-----SPGCGSPSPSPSNST-TTPPPASAP-PTTP---SSPPPTPTPTSP 48
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 MATIPVQPVVSNPPVTSPPPLNNATSPATPPVTSPLPPSAPPNRPAPPPVPTSP 60
QY 49 PPSRSPT-----SAPPSPPTPTSPGPPPLPQP-----SPP---ATTP 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 PFVANGAPPPLPKAPPESSSPFPQPVIPSPSPSTSPFPVPSPPSAPPALVPLP 120
QY 87 GSPAP--VTP-----TRNPPPSV-----PQPSN-PREGGSPRPPSP-- 124
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 SSPPPASVPPPRSPSPPLVRSPPSVRIQSPPPPSDRPQSPPPSPSPSPSERP 180
QY 125 --SPPSSDGLSTGVVGAIGVALLVITLICLCCKKRRDEADYVPPPPPPG 182
||||| |
181 TQSPSPSPS-----ERPTQSPSPSPSP-- 202
QY 183 KAGPGYGCQQQWQQNATPPSDHVTSILPPPK--APSPRPPPPPPPPPPFMS----- 234
||||| |
203 -----SPPSRPQSPPPPPPEDIKQPPRSPNPPPPPTFFSPRSP 244
QY 235 -----SSGGSDYSDRVPVLP-- 250

```

Db 245 EILVPGSNPNPSONNPTLRPELDAPNSTNNNGIGTGAIVWGIVSVAVALVFTLFGIFWCLR 304

QY 251 -----SPGLV-----LGFST 262

Db 305 KREKRLSAYSGDVYPSPMSSTARSASFRRQSSAPVAGSKRSQSGSGGLGNSKAL 364

QY 263 FYEELARATNGFSEANLLIQGGGFGYVHKVLPSPGKEVAVKQKVGSGGGEREFQAEVEI 322

Db 365 FSYEELVKATNGFSQENLLGEGFGCVYKGLPDRVAVKQKIGGGQGDREKAEVET 424

QY 323 ISRVHHRHLVSVGVCIAGAKRLLYVEFPNNNLEHLHCEGPTMEWSTRLKIALGSAK 382

Db 425 LSRIRHRLVSVGVHCISGRRLLYVSNNDLYEHLHGE-KSVLDNATRVKIAAGAR 483

QY 383 GLSYLHEDCNPKIIRHDIKASNLIDFKFEAKVADFLAKIASDTHVSTVWGTFGL 442

Db 484 GLAYLHEDCHPRIIRHDIKSSNLLIEDFARVSDGLARLALDCNTHITTRVIGTFGYM 543

QY 443 APEYASGKLTSEKSDVSEFVGVLELITGRPRVDANNVYVDSLVVDWARPLLRASEQGD 502

Db 544 APEYASGKLTSEKSDVSEFVGVLELITGRPKVDTSQPLGDESLSVWARPLISHATEEE 603

QY 503 FEGLDADAKNNGYDREEMARMVACAAACVHRSARRSPMSQIVRALEGVNSVSLDNEGR 562

Db 604 FUSLDADKLGNGYVSEMFRTIEMAGACVHRHATKSPRMGQIVRAFE-SLAEDLTNGMR 662

QY 563 PQGSNVYSYSGGSTDYDSSQYNEDMKKFRKMGALGICEYN 601

Db 663 LGSEV-----FNSAQSAELRFRMAFGSQNYS 692

RESULT 13

Q9FFW5 PRELIMINARY; PRT; 681 AA.

AC Q9FFW5; (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1];

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,

RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,

RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,

RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,

RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,

RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,

RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,

RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [3];

RP SEQUENCE FROM N.A.

RA Shin P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,

RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,

RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,

RA	Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA	Seki M., Souttack A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA	Ecker J.R.;
KT	"Arabidopsis cDNA clones";
KL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB05231; BAB10145.1; -
DR	EMBL; AF424623; AAL11616.1; -
DR	EMBL; AY075681; AAL77688.1; -
DR	InterPro; IPR000719; Euk_pkinase.
DR	InterPro; IPR003882; Pistil_extensin.
DR	InterPro; IPR002965; P_rich_extensn.
DR	InterPro; IPR004040; Ser_thr_pkinase.
DR	InterPro; IPR001245; Tyr_pkinase.
DR	Pfam; PF00069; pkinase; 1.
DR	PRINTS; PR01217; PRICHEXTENSIN.
DR	PRINTS; PR01218; PSTLEXSENSIN.
DR	PRODOM; PD000001; Euk_pkinase; 1.
DR	SMART; SM00221; STKc; 1.
DR	SMART; SM00220; S_TKc; 1.
DR	SMART; SM00219; TYRKc; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	ATP-binding; Kinase; Transference.
SQ	SEQUENCE 681 AA; 72389 MW; F64DAA12470E73F9 CRC64;
Query Match	42.28; Score 1456; DB 10; Length 581;
Best Local Similarity	46.48; Pred. No. 2.5e-85;
Matches 306; Conservative 87; Mismatches 158; Indels 108; Gaps 14;	
QY	1 MSSAPSEGTGSPSPSPSSNTTTTPPASAPPTTSPSSPPSTIPTSPSSSRSTPSAPP 60
Db	: : : : : : : : : : :
52	VSSSPPEVSSP-PPSSSPSPSPVILSPPTVASSPPPVVIASPPSPPTPATFPAPP 110
QY	61 -----PSPPTPSTPGSP-----PPLQPSPPAFTTSGS 88
Db	: : : : : : : : : :
111	QTVPSPPPPDAASPAPTITNPKPKSPSPGGSTSPGPETSPKPSPSTPTPTTTS 170
QY	89 PPAP-----VTPTRNP-----PSPVGP-----PSNPSRGSGRPPSSSP 127
Db	: : : : : : : : : :
171	PPPPATSAFPSNPTDPSTLAPPTLPVPREKPIAKPTGASNNGNTLPSS---- 226
QY	128 SPSDGLSTG--VVVGIAIGVALLIVTLICLCCKKRREDDAYVPPPPPGPKAG 185
Db	: : : : : : : : : :
227	SPGKSEVTGGIAGVIGVLSELTFVMGVMTXKRXKDPGFVGYTMP-----SAYS 262
QY	186 GPYGECQQQRQNAPPSPDHVTVSLPFPKAFSPRPQPPPPPMSSSGSDYSDRP 245
Db	: : : : : : : : : :
283	SPGGSQVLFNSRSSAPPK-----MRSHSGDY---- 310
QY	246 VLPSPGPLVLGFSKSTFTVEELARATNGFSEANLLSQGGFGVHKGLVPSGEVAVKQL 305
Db	: : : : : : : : : : :
311	MYASSDSGMYSN-QRSWFSDYDELQVTSGESEKNLLGEGFGCVYKGVLSDGREVAVKQL 349
QY	306 KVSGGGGEREQAEVELISRVHRLVSLVGYCIAGAKRLLVTFEFPNNNLLELHGGR 365
Db	: : : : : : : : : : :
370	KIGSGGEREFKAEEVILSRVHHRLVTLVGYCTSEQHRLLVTDYVENNTLVHLLHAPGR 429
QY	366 PTMWSLRKLIALGSKLSYLHDCNPKLIHRDIKASNILIDEKFPAKVADFLAKIAS 425
Db	: : : : : : : : : :
430	PWTWEUFRVAAGAARGIAYLHEDCHPRIIHRIKSNIIILDNFSFAELVADEFLAKIAQ 489
QY	426 --DTNTHVSTRVMGTFFGYLAPEAASGLKEKSVDTSFGVVLLELTIGRPVDANNVVD 483
Db	: : : : : : : : : :
490	ELDINTFVSTRVMGTFTGYMAPEVATSGKLEKADVTSYGVILLELTIGRPVDTSQLGD 549
QY	484 DSVLDWARPLLNRASEQGDFEGLADAKMNNGYDFEEMARMVACAACVRSAARRRRPM 543
Db	: : : : : : : : : : :
550	ESLVEWARPLLGAIEENEEDELVDPLIGNFLTGEMFRVYEAAACVRSAARKPKMSQ 609
QY	544 IVRALGNVSLDLNEMGRPQSNVTSYSGGSTDYDSSQYNEDMKFKRMALGTQETNA 602

